Diffusion of Electronic Health Records

Nøhr, Christian; Andersen, Stig Kjær; Bernstein, Knut; Bruun-Rasmussen, Morten; Vingtoft, Søren

Published in:
Medinfo 2007. Proceedings of the 12th World Congress on Health (Medical) Informatics

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

MEDINFO 2007

Studies in Health Technology and Informatics

This book series was started in 1990 to promote research conducted under the auspices of the EC programmes' Advanced Informatics in Medicine (AIM) and Biomedical and Health Research (BHR) bioengineering branch. A driving aspect of international health informatics is that telecommunication technology, rehabilitative technology, intelligent home technology and many other components are moving together and form one integrated world of information and communication media. The complete series has been accepted in Medline. Volumes from 2005 onwards are available online.

Volume 129

Recently published in this series

ISSN 0926-9630

MEDINFO 2007

Proceedings of the 12th World Congress on Health (Medical) Informatics

Building Sustainable Health Systems

Part 2

Edited by

Klaus A. Kuhn

University Medical Center, Technische Universität München, Germany

James R. Warren

Department of Computer Science, University of Auckland, New Zealand

and

Tze-Yun Leong

School of Computing, National University of Singapore, Singapore

IOS Press

Amsterdam • Berlin • Oxford • Tokyo • Washington, DC

ISBN 978-1-58603-774-1

Publisher
IOS Press
Nieuwe Hemweg 6B
1013 BG Amsterdam
Netherlands
fax: +31 20 687 0019
e-mail: order@iospress.nl

Distributor in the UK and Ireland
Gazelle Books Services Ltd.
White Cross Mills
Hightown
Lancaster LA1 4XS
United Kingdom
fax: +44 1524 63232
e-mail: sales@gazellebooks.co.uk

Distributor in the USA and Canada
IOS Press, Inc.
4502 Rachael Manor Drive
Fairfax, VA 22032
USA
fax: +1 703 323 3668
e-mail: iosbooks@iospress.com

PRINTED IN THE NETHERLANDS

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Table of Contents

Part 2

Chapter 5. Decision Support and Workflow 827

Chapter 5.

Decision Support and Workflow

This page intentionally left blank

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

How Updating Textual Clinical Practice Guidelines Impacts Clinical Decision Support Systems: a Case Study with Bladder Cancer Management

Jacques Bouaud[a], Brigitte Séroussi[b], Ambre Brizon[a], Thibault Culty[c], France Mentré[d], Vincent Ravery[c]

[a]*AP-HP, DSI, STIM, Paris, France; INSERM, UMR_S 872, eq. 20, Paris, France*
[b]*Université Paris 6, UFR de Médecine, Paris, France; AP-HP, Hôpital Tenon, Département de Santé Publique, Paris, France*
[c]*Université Paris 7, UFR de Médecine, Paris, France; AP-HP, Hôpital Bichat, Service d'Urologie, Paris, France*
[d]*INSERM, U738, Paris, France; Université Paris 7, UFR de Médecine, Paris, France; AP-HP, Hôpital Bichat, Paris, France*

Abstract

Guideline-based clinical decision support systems (CDSSs) can be effective in increasing physician compliance with recommendations. However, the ever growing pace at which medical knowledge is produced requires that clinical practice guidelines (CPGs) be updated regularly. It is therefore mandatory that CDSSs be revised accordingly. The French Association for Urology publishes CPGs on bladder cancer management every 2 years. We studied the impact of the 2004 revision of these guidelines, with respect to the 2002 version with a CDSS, UroDoc. We proposed a typology of knowledge base modifications resulting from the update of CPGs making the difference between practice, clinical conditions and recommendations refinement as opposed to new practice and new recommendations. The number of formalized recommendations increased from 577 in 2002 to 1,081 in 2004. We evaluated the two versions of UroDoc on a randomized sample of patient records. A single new practice that modifies a decision taken in 49% of all recorded decisions leads to a fall from 67% to 46% of the compliance rate of decisions.

Keywords:

clinical decision support system, practice guideline modelling, knowledge base revision, bladder cancer

Introduction

Clinical practice guidelines (CPGs) are developed and disseminated to promote interventions of proven benefits. CPGs are produced by health professional societies or national health agencies relying on evidence-based medicine principles. They are made of structured, explicit, motivated, textual statements for managing patients according to the state of the art. However, in the "knowledge crisis" era, the amount of new medical knowledge endlessly increases [1] and the meaning of "state of the art" is a relative and evolving notion. As a consequence, CPGs are involved in a life-cycle and must be regularly updated. This process is handled by CPG providers when new research results modify current best practices. It takes at least one year to publish an updated version of CPGs.

Despite their broad dissemination, there is still considerable variation in the effectiveness of guidelines to change physicians' behavior. Many studies have shown that clinical decision support systems (CDSSs) can be effective in increasing physician compliance with recommendations [2, 3]. Defined as any software in which characteristics of individual patients are matched to a computerized knowledge base (KB) for the purpose of generating patient-specific assessments or recommendations that are presented to clinicians for consideration, CDSSs rely on formalized models of textual guidelines. Hence, as these knowledge sources are revised, it is mandatory that the delivery instrument be revised accordingly in a continuous knowledge management process. However, encoding original guideline texts into a computerized format is still an issue many authors are addressing [4, 5, 6, 7]. The management of multiple, successive, versions of a guideline is a problem CDSS developers are aware of for a long time.

In this paper, we have studied the impact of the publication of a new version of a guideline on a CDSS. We used the bladder cancer guidelines developed by the French Association for Urology ("Association Française d'Urologie" or AFU) in its revisions of years 2002 and 2004. For each textual versions, we built the corresponding KB of a CDSS, named UroDoc, and obtained 2 versions of the system: UroDoc-2002 and UroDoc-2004. After a study of both versions of textual CPGs, we first compared the two corresponding KBs. Then, we assessed how actual practices were impacted by the publication of the 2004 CPGs. We carried out a retrospective study on a randomized sample of medical records from patients hospitalized in 2004 at the Department of Urology of the Bichat-Claude Bernard Hospital, Paris, France.

Selected for best paper award.

Material and method

Textual bladder cancer guidelines

With more than 10,000 cases yearly, bladder cancer is the forth most common cancer among French men and the eighth in women. Bladder cancer represents a range of disease from relatively benign surface tumors to highly malignant life threatening carcinomas which require radical treatment. If approximately 70% of patients with newly diagnosed bladder cancer will present with superficial bladder tumors, most of them will eventually have recurrent cancer, and 10 to 20% will progress to muscle-invasive or metastatic disease. Treatment of superficial tumors is mainly based on transurethral resection (TUR). Adjuvant bladder instillation of mitomycin C (MMC) or BCG may be indicated. A strict and prolonged follow-up is mandatory to detect recurrence or progression. The standard treatment for patients with invasion remains cystectomy. However, surgical cure rates are only in the range of 10 to 30% in locally advanced bladder cancers. At present, standard therapy is still missing for numerous patient clinical profiles, hence making bladder cancer management a very active area of research.

Founded in 1896, the French Association for Urology represents and supports most of the French urologists. An expert group, the AFU Oncology Committee, publishes CPGs updated every two years to disseminate new management standards and improve healthcare quality and survival rates. Bladder cancer CPG revisions have been regularly published in 1998, 2002 and 2004. We studied bladder cancer CPGs evolution between 2002 and 2004.

Documentary-based decision support paradigm

Numerous studies have shown that the sole dissemination of textual CPGs is inefficient to modify physician behavior. On the contrary, computer-based reminders, intervening at decision time, to automatically provide patient-specific recommendations from coded patient data, seem to be the most efficient means to influence healthcare professionals in the adoption of CPGs [2, 3].

However, the success of reminder-based CDSSs is not warranted [3]. Relevance and accuracy of the advices automatically provided are often weak: Judge *et al.* [8] reported that only a little half of alerts were directly relevant to medication order when using a CDSS, Van der Sijs *et al.* [9] explained that this may lead clinicians to override 49 to 96% of the alerts. In addition, the lack of flexibility of automated approaches is often criticized: if reducing the complexity of a given patient to a set of *data* is a necessary step to feed the electronic medical record, the medical reasoning process involved in medical decision making requires a flexible and contextual interpretation of involved medical *notions*.

That is why other approaches, where guideline knowledge is structured in a way a user could retrieve patient-specific recommendations more easily than within texts, have been proposed to provide physicians with guidance [10]. The OncoDoc system, applied to breast cancer therapeutic management, has been developed according to these principles.

We developed a comparable system, UroDoc, to implement the bladder cancer CPGs of year 2002 (UroDoc-2002) and 2004 (UroDoc-2004). In each case, the KB is formally structured as a decision tree. The system need not be automatically executed from an electronic medical record. The physician can browse the KB hence controlling how each patient criterion is instantiated; she has the opportunity to assess the patient clinical states that do not need to be encoded and can be available on different supports. Thus, she identifies the path of the decision tree that corresponds to her patient and obtains in a flexible way the relevant patient-specific recommended therapy.

Modelling textual CPGs into structured decision trees

The two textual CPGs were studied in chronological order. For each CPG, the modeling work has been handled independently, *i.e.* we didn't try to build the version $d + 1$ of the CPG from its differences with version d. Following the approach to guideline implementation proposed by Shiffman *et al* [4], we first proceeded to the atomization step to extract single concepts from the recommendation's natural language text (*microscopic hematuria, grade, intravesical therapy*, etc.). Then, we carried on the deabstraction step to adjust the level of generality at which decision variables or actions were described in the CPGs to permit the operationalization in a clinical setting ("*early recurrence*" has been translated in "*recurrence at the same stage as the initial tumor or at a more advanced stage within 3 months*"). Then the disambiguation step tried to establish a single semantic interpretation for the recommendation statement (how "*biopsies may be performed on suspicious area*" should be interpreted?). The most important step was the verification of completeness to assure that both implicit decision criteria and implicit modalities of identified criteria have been explicited as decision variables and actions, and that there was a treatment (either a recommendation or a professional agreement depending on the availability of evidence) proposed for all logically possible combinations of condition states meaning that the whole set of clinical situations described by the complete expansion of the decision tree is addressed. Two different decision trees have thus been built from the two CPGs.

Comparison of the two versions of CPGs

Theoretical study

We first compared the original two textual documents from a macroscopic point of view, in terms of title, authors, general outlook, logical structure, length of comparable sections. Then we compared the two KBs built from the texts. To evaluate the evolution of CPGs between 2002 and 2004, we proposed a typology of the changes that may be found between consecutive versions of documents. CPGs can be considered as a set of recommendations $\{R_i\}_i$. Each recommendation R_i is characterized by a pair (S_i, T_i) with $S_i \rightarrow T_i$, denoting that a treatment plan T_i is recommended in the clinical situation S_i. As usual, a clinical situation is defined by a set of instantiated decision variables, $S_i =$

$\{c_{i,j}\}_j$, and a treatment plan is defined by an ordered sequence of therapeutic actions, $T_i = \{t_{i,j}\}_j$.

The set of recommendations has to be dated, thus we have to compare recommendations $\{R^d_i\}_i$ in their *d*-version to their *d+1*-version $\{R^{d+1}{}_i\}_i$. For simplicity reasons, we have dropped indices in the rest of the section. We consider that two clinical situations S^d and S^{d+1} are comparable, noted $S^d \cong S^{d+1}$, when both situations correspond to the same clinical profile in terms of therapeutic history, staging and grading of the tumor. We consider that two treatment plans T^d and T^{d+1} are comparable, noted $T^d \cong T^{d+1}$, when both treatment plans correspond to the same sequence of care in terms of therapeutic modalities (surgery, immunotherapy, chemotherapy, radiation therapy) and, in case of a surgery, when it applies to the same anatomical part. The set of all possible modifications that can occur in the update of a given CPG is illustrated by figure 1.

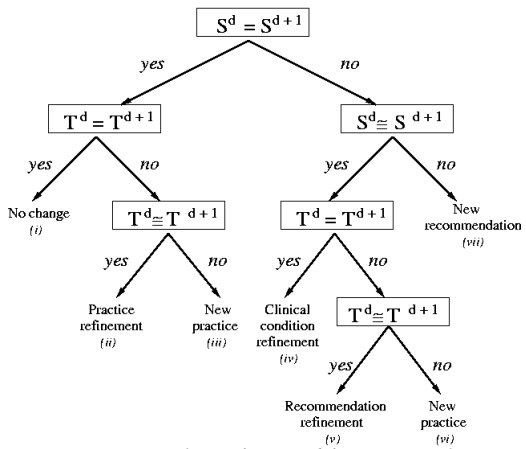


Figure 1 - Typology of KB modifications resulting from CPGs updating

Four main situations can be observed in the update process:

1. No change: an identical clinical situation lead to an identical treatment plan *(i)*.
2. Refinement of an existing recommendation. The new recommendation shares some common parts with the former one. Refinement may concern: the description of the treatment plan but not the situation *(ii)*, only the description of the clinical situation *(iv)*, or both descriptions *(v)*.
3. New practice. A totally new therapy (non comparable) appeared in an already identified clinical situation *(iii)*, or in its corresponding refined expression *(vi)*,
4. New recommendation: extending the CPG coverage, a new clinical situation is identified with its corresponding therapy leading to a new recommendation *(vii)*.

When comparing the two KBs, decision variables, single therapeutic actions, and treatment plans were analyzed literally to determine those which were common, *i.e.,* used in both KBs, those which were 2002-specific, *i.e.,* obsolete in 2004, and those which were 2004-specific, *i.e.,* new in 2004. In this first step, we used the identity relationship between objects (=). The same process was applied at a more abstract level where decision variables and single therapeutic actions were substituted by their upper concept in a kind of ontological reasoning. The relationship used for comparison was the comparability (≅). For instance, "chemo-III" and "chemo-IV" correspond to two therapeutic actions, both abstracted as "adjuvant-chemotherapy", and therefore considered comparable.

Practical analysis

We studied the impact of the publication of the new version of the bladder cancer management guideline in June 2004 upon the clinical practice of the Department of Urology of the Bichat-Claude Bernard hospital in Paris, France. We collected a random sample of patients diagnosed with bladder cancer during the year 2004. For each therapeutic decision following the diagnostic of bladder cancer, we used both systems UroDoc-2002 and UroDoc-2004 to get the recommendations of 2002-guidelines and of 2004-guideline. These two therapeutic recommendations have been compared to the treatment stored in the medical record and actually received by the patient.

Results

Comparison of textual CPGs

Both 2002 and 2004 CPGs have the same title anc coverage. They have the same number of authors (15), 80% of them participated to both CPG developments. Both guidelines are published as textual double-column documents. 2002 CPGs are shorter (13 pages) than 2004 CPGs (40 pages). General outlook is comparable (Table 1), starting with a description of the staging system (tumor, node, metastasis, TNM), followed by 3 sections describing diagnostic indications, therapeutic guidelines and follow-up protocols. However, some sections especially those describing the management of patients with positive nodes and metastatic disease are more elaborated in the 2004 version. Besides, a new sub-section specifying pathologic staging modalities has been added in the 2004 version.

Table 1 - CPG structure and volume in columns (half-pages)

	Structure	*Number of columns*	
		2002	*2004*
0.	**Introduction**	**1**	**1**
I.	**TNM staging system**	**1**	**1**
II.	**Diagnostic indications**	**1.5**	**5**
1.	*Clinical presentation*	*0.5*	*1*
2.	*Imaging tests*	*0.5*	*1.5*
3.	*Urine cytology*	*0.25*	*0.5*
4.	*Cystoscopy*	*0.25*	*1*
5.	*Pathologic staging*	*n/a*	*1*
III.	**Therapeutic guidelines**	**8**	**21.5**
1.	*Therapies*	*5*	*6*
2.	*Management principles*	*3*	*15.5*
2.1.	*Superficial tumors*	*1.5*	*2.5*
2.2.	*Invasive tumors (N0M0)*	*1*	*5*
2.3.	*Positive nodes (N+M0)*	*0.25*	*6*
2.4.	*Metastatic disease (M+)*	*0.25*	*2*

Comparison of structured KBs

Table 2 reports the formal comparison of the two KBs. It should be noticed that at the literal level, the 2004 version, although sharing common elements with the 2002 version, introduced many new ones and left apart some. Although reducing the importance of the effect, the abstraction step confirms the trend. When considering decision trees, the 2002 version identifies 366 different clinical situations whereas the 2004 version identifies 584 clinical situations. The average path length is 10 in 2002 but 11 in 2004. There are 577 recommendations in 2002 and 1,081 in 2004.

Table 2 - Comparisons of decision variables (DVs), therapeutic actions (TAs), recommendations (Recos), and of their abstracted correspondents in 2002 and 2004 KBs

	2002 specific	*Common*	*2004 specific*
DVs	8	32	23
Abstracted DVs	6	28	13
TAs	9	27	22
Abstracted TAs	4	17	5
Recos	530	47	1,034
Abstracted Recos	339	238/358	723

When comparing the two sets of recommendations, $\{R^{2002}{}_i\}_i$ and $\{R^{2004}{}_i\}_i$, we adopted two different points of view. From the 2002 point of view, each recommendation has been tested against the 2004-recommendation set to automatically determine its status: whether it remained unchanged, it was refined, it has evolved towards new practices, or it became obsolete (*i.e.,* it has no more equivalent in 2004). The same algorithm was applied for each 2004 recommendations as compared to 2002 recommendations. Table 3 reports these results and figures 2 and 3 illustrate the percentages of the different subsets.

Table 3 - Distribution of recommendation evolution status between 2002 and 2004 versions of the KB

	2002 KB	*2004 KB*
Obsolete recommendations	339	—
No change	47	47
Refinement	74	91
New practices	117	220
New recommendations	—	723
Total	577	1,081

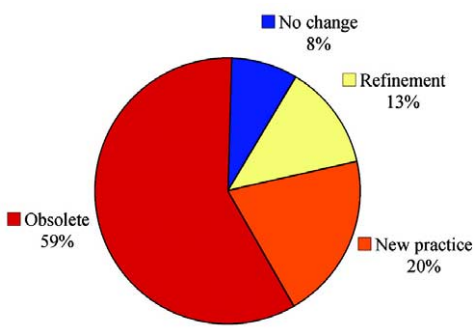


Figure 2 - Status of 2002 recommendations (n=577)

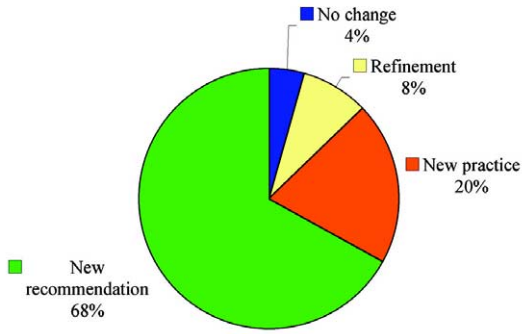


Figure 3 - Status of 2004 recommendations (n=1,081)

The overlapping of the 2004 KB on the 2002 KB is quite limited. The overlapped part (i.e., no change, refinement, new practice) is made of 238 recommendations (41%) in 2002 which projection is made of 358 recommendations (33%) in 2004. From this analysis, 339 recommendations of 2002 (59%) became obsolete in 2004 and 723 recommendations of 2004 (68%) may be considered as new.

Impact of CPG update on clinical practices

We analyzed medical records of a randomized sample of 45 patients leading to a total of 106 therapeutic decisions. Only 97 decisions, corresponding to 38 patients, were retained for the survey: 9 decisions were ruled out either because the patients were long lost (6), or decisions also involved the management of prostate cancer (3). Each therapeutic decision has been retrospectively matched to the 2002 CPGs by using UroDoc-2002 and to the 2004 CPGs by using UroDoc-2004. With respect to the 2002 CPGs, we found that 65 decisions, *i.e.,* 67% of the cases, were compliant with the CPGs. When the same therapeutic decisions were evaluated using UroDoc-2004, we found that 45 decisions, *i.e.,* 46% of the cases, complied with the 2004 CPGs. Among the non-compliant decisions, only a few cases (5 as compared to 2002 CPGs and 4 as compared to 2004 CPGs) can be explained by specific patient conditions (age) or "acceptable" physician personal choice. Table 4 synthesizes the results of the analysis.

Table 4 - Practical compliance with 2002 and 2004 CPGs

(n=97)	*# non compliant*	*# explainable*	*Compliance*
2002	32	5	67%
2004	52	4	46%

Discussion and conclusion

The 2004 evolution of bladder cancer CPGs essentially concerned the extension of CPGs coverage and thus increased the number of clinical situations for which recommendations are provided. This is particularly true for the management of patients with positive nodes and metastatic disease. This is also the case for the control of BCG side-effects which is not mentioned in 2002 while very much detailed in 2004. A particular new practice has to be

noticed: TUR which was recommended as such in 2002 has to be followed by an early post surgical endovesical instillation (EPSEI) of MMC in 2004. When analyzing the compliance of therapeutic decisions with respect to Uro-Doc-2002 and UroDoc-2004, we got the results summarized in table 5.

Table 5 - Evolution of decision compliance with 2002 and 2004 CPGs

	2002 compl.	*2002 non compl.*	*Total*
2004 compl.	41	4	45
2004 non compl.	24	28	52
Total	65	32	97

The 24 decisions that are 2002-compliant, but not 2004-compliant, are made of 21 TUR with no EPSEI, 1 decision of bladder instillation of BCG with no endoscopic evaluation, and 2 incomplete surgeries as compared to 2004 recommendations. When studying the clinical practices of the Department of Urology of the Bichat-Claude Bernard hospital, it appears that 49% of the therapeutic decisions for bladder cancer correspond to TUR decisions (figure 5). Since 2004 recommendations modified the TUR decision, this illustrates the fact that, although they were published, new 2004 CPGs were not applied in routine care. The side-effect of this delay in CPG implementation is a fall from 67 to 46% of the compliance rate with CPGs.

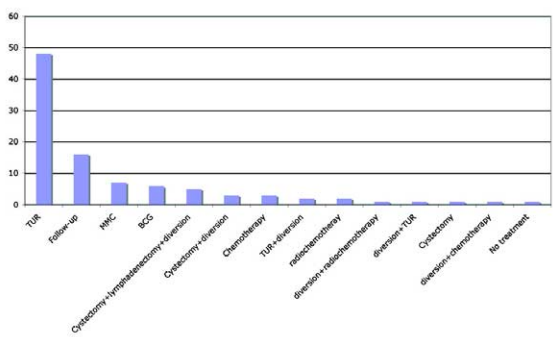


Figure 5 - Distribution of actual therapeutic decisions (n=97)

Besides, the formal comparison of KBs showed that the number of recommendations increased by a factor of 2. It is remarkable that more than half of the 2002 recommendations (339/577) have no formal equivalent in the 2004 KB since this is not consistent with CPG content. The 2004 version is supposed to cover every aspects of the 2002 version and the 339 "unmatched" recommendations should have their equivalents at the semantic level in the 723 "identified as new" 2004 recommendations. The reason is that every 2002 recommendation which contains at least one of the 8 2002-specific variables never matches any 2004 recommendation. This emphasizes the fact that in 2002 and 2004, some similar notions expressed differently have been modeled differently. As a conclusion, KB revisions following CPG updating could be improved if the new CPG version respects the terminology and point of view used in the former version. This could be achieved by explicitly using, when possible, standard terminologies to describe guideline-covered specific clinical situations in CPGs.

References

[1] Wyatt J. Use and sources of medical knowledge. Lancet 1991;338(8779):1368–73.

[2] Garg AX, Adhikari NKJ, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyenne J, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. JAMA 2005;293(10):1223–38.

[3] Shiffman RN, Liaw Y, Brandt CA, and Corb GJ. Computer-based guideline implementation systems: a systematic review of functionality and effectiveness. JAMIA 1999;6(2):104–14.

[4] Shiffman RN, Michel G, Essaihi A, and Thornquist E. Bridging the guideline implementation gap: a systematic, document-centered approach to guideline implementation. J Am Med Inform Assoc 2004;11(5):418–26.

[5] Svatek V and Ruzicka M. Step-by-step formalisation of medical guideline content. Int J Med Inf 2003; 70(2- 3):329–35.

[6] Votruba P, Miksch S, Seyfang A, and Kosara R. Tracing the formalization steps of textual guidelines. In: Kaiser K, Miksch S, and Tu S, eds, Computer-based Support for Clinical Guidelines and Protocols. Proc Symposium on Computerized Guidelines and Protocols (CGP 2004), (vol 101) of Studies in Health Technology and Informatics, Amsterdam. IOS Press, 2004:172–6.

[7] Shahar Y, Young O, Shalom E, Galperin M, Mayaffit A, Moskovitch R, et al. A framework for a distributed, hybrid, multiple-ontology clinical-guideline library, and automated guideline-support tools. Journal of Biomedical Informatics 2004;37(5):325–44.

[8] Judge J, Field TS, DeFlorio M, Laprino J, Auger J, Rochon P, et al. Prescribers' response to alerts during medication ordering in the long term care setting. J Am Med Inform Assoc 2006;13(4):385–90.

[9] van der Sijs H, Aarts J, Vulto A, and Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc 2006;13(2):138–47.

[10] Séroussi B, Bouaud J, and Antoine EC. OncoDoc, a successful experiment of computer-supported guideline development and implementation in the treatment of breast cancer. Artif Intell Med 2001;22(1):43–64.

Address for correspondence

Jacques Bouaud
STIM/DSI/AP-HP
3, avenue Victoria, F-75004 Paris, France
jacques.bouaud@sap.aphp.fr

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Improving Compliance to Guidelines through Workflow Technology: Implementation and Results in a Stroke Unit

Silvia Panzarasa[a], Silvana Quaglini[b], Giuseppe Micieli[c], Simona Marcheselli[d], Mauro Pessina[e], Corrado Pernice[e], Anna Cavallini[d], Mario Stefanelli[b]

[a] *Consorzio di Bioingegneria e Informatica Medica, Pavia, Italy*
[b] *Dip. Informatica e Sistemistica, Università di Pavia, Italy*
[c] *Neurology and Stroke Unit, IRCCS Istituto Clinico Humanitas Rozzano (MI), Italy*
[d] *IRCCS Fondazione Mondino, Stroke Unit, Pavia, Italy*
[e] *TSD Project, Milan, Italy*

Abstract

This work describes the results of the implementation of a workflow management system integrated into the electronic clinical chart of a Stroke Unit. The workflow logic is based on the rules provided by the SPREAD guidelines for stroke management. In this way, the already existing clinical chart has been transformed into an evidence-based, real-time decision support system, meanwhile maintaining the same look the users were familiar with. Since the final aim of the work was to improve evidence-based behavior and detect possible organizational bottlenecks, non-compliance to the clinical practice guidelines, before and after the system introduction, have been analyzed, as well as the accuracy of the clinical chart compilation, some care process variables, and system usability. Results show that the system enhances the clinical practice without boring users. Moreover, non-compliance analysis gives rise to ideas for further improvement.

Keywords:

decision support systems, practice guideline, information systems, workflow management, stroke

Introduction

Insufficient or erroneous information at point-of-care, both for diagnosis and treatment, is a frequent and significant cause of medical error, as well as communication problems among healthcare operators and the different healthcare organizations, where patients' data are spread. Prevention of possible errors during a patient's management is a critical issue. Recent studies have highlighted the dramatic dimension of the problem, from the well-known report of the US Institute of Medicine [1], assessing between 44,000 and 98,000 deaths per year as due to medical errors in American hospitals, to an increasing number of country-based studies reporting similar figures.

Evidence-based clinical practice guidelines (GLs) have been widely promoted as a way of improving health outcomes. In the stroke management area, in particular, the study described in [2] demonstrates an association between stroke outcome and compliance with GLs. In current healthcare systems, however, scientific knowledge about best care is not applied systematically or expeditiously to clinical practice. GLs usually capture both literature-based and practice-based evidence into a textual format, which can be easily diffused but uneasily used in daily work. Thus there is a great effort to disseminate them in computer-interpretable representations, more suitable for individual clinical decision support [3]. Contemporary, there is increasing need of smooth integration of GLs into the existing hospital information systems.

The current workflow technology [4] seems to offer a convenient solution to build a cooperative system in which the activities of a care providers' team can be coordinated within a process properly designed on the basis of available best medical knowledge. This paper presents an approach to the design, implementation and evaluation of an evidence-based workflow management system (WfMS), from here on called Careflow management system (CfMS). On the basis of a general methodology, we describe a CfMS implementation in the area of Stroke management, particularly focusing on user interaction, presentation of GL suggestions and management of non-compliance to GL. We show how this technology, especially in fields where acute patients treatment is critical, helps in timely collecting the correct and most useful data and avoiding errors. The test-bed for the proposed methodological and technological solutions is the Stroke Unit (SU) in Pavia, where we implemented the system based on the SPREAD guidelines (Stroke Prevention and Educational Awareness Diffusion) [5].

The evaluation methodology

Outcomes of the CfMS have been tested following the steps below, described in detail in the subsequent sections:

1. analysis and update of the Computerized Clinical Chart (CCC) that physicians were using since many years in the SU, with the aim of ensuring that all the necessary data for testing the care process could be stored;

2. utilization of the updated CCC, without any decision support utility: during this phase, compliance to GLs was exclusively based on physicians' personal knowledge and on hospital organization;
3. development of the decision support tool, i.e. the CfMS, and its integration with the CCC;
4. utilization of the integrated system for 8 months (from its installation until November 2006). During this phase, compliance could have been affected by system reminders; and
5. analysis of collected data.

Updating the stroke unit data model

The Stroke Unit was already equipped with a CCC implemented some years ago, with the commercial tool Wincare® (by TSD-Projects). System-user interaction is based on so-called *Events*: an event is a single data entry form of the clinical chart, and it may be activated by users with competent role (controlled access). Before starting the project we analyzed all the events and their associated graphical user interface, with the aim of checking the presence of the data required for both properly executing the decision support system and evaluating GL compliance. As a result, the data model has been updated both by increasing the number of data input fields and by changing the nature of some existing data, mainly shifting from free text to encoded data. In fact, data-entry was mainly limited to free-text forms, because the initial purpose of the CCC was only to limit paper-based communication and to improve patients' data retrieval. The advantage was that information was ready to be printed *as-it-was* for summaries, discharge letters, etc.. Moreover, such CCC gave users the flexibility to represent patient's condition in the desired order and granularity. On the other hand, integration with a decision support system was not possible because of almost complete lack of structured data. In other words, free-text information stored in the CCC didn't allow a proper interpretation of the rules embedded in the GL: natural language processing is a hot research field, but results are not mature enough to guarantee full recognition of crucial data in critical patients' management [6].

Eventually, a deep examination of SPREAD GLs allowed determining the minimum data set required for implementing all the recommendations. These data have been encoded and, when possible, standard terminologies such as ICD9-CM for diseases, have been used. To maintain flexibility, in addition to encoded items, the interface still allows entering free text notes, useful to input additional details but not essential to decision support purposes.

Careflow implementation and integration

After the analysis and the consequent update of the data model, we implemented the CfMS, based on all the GL recommendations from SPREAD chapters 9 and 10, related to diagnosis and treatment of the stroke acute phase, with some site-specifications decided by the neurologists involved. Technology used to implement the whole system and to integrate it with Wincare® is based on DBMS Oracle™, Oracle Workflow™, Pl/SQL and Visual Basics (see [7] for details). Briefly, the CfMS manages the execution of the care process for a particular patient, reading his/her clinical data, verifying the GL rules and creating specific recommendations. These latter, translated in to-do lists, messages and alerts, are communicated to the right professional role via the end-user application.

Since the users were satisfied with their CCC, our choice was not to create an additional and specific interface, but to integrate all the needed functionalities within the existing CCC, making it more "dynamic" and "smart" thanks to the CfMS execution. Our philosophy was that users must perceive the new system just as an update, with some new functionalities. To obtain this goal, a middleware layer for the data sharing has been developed in order to keep the two systems (Wincare® and CfMS) independent, while granting communication. In particular, whenever there is a new data entry, the CCC transfers these data into a buffer database, usable by the decision support system and designed on the basis of the minimum data set required by the GL. Similarly, when the CfMS generates a suggestion, or a new to-do list, they are put into special tables of the same database, and Wincare® can read and show them through its interface.

Relevant functionality

A picture of the updated system interface is shown in Figure 1.

Facilitating data input

First of all, data relevant for GL rules interpretation appear as yellow input fields, in order to encourage a complete editing of the essential fields of the CCC.

Smart graphical user interface

With respect to the former system, the list of the CCC events has been transformed into a "to-do" list, i.e. events are now listed in a patient-specific, dynamic way, taking into account the patient medical condition and the time spent since his/ her admission. Different icons and colors are used to mean that an event is completed (icon with a green happy smiley), it is currently being done (orange thoughtful smiley), it has not been executed due to an exception (red sad smiley), it has been filled on the user initiative (Wincare® icon), or it still has to be done (no icon). Degree of the scientific evidence of recommendation is shown, as in the SPREAD book, by letters A-D. According to the physician needs, special attention has been put in avoiding over-information: particularly in acute situations, only events necessary to face the urgency are listed. For example, if a patient is potentially eligible for r-tPA treatment (a thrombolitic drug very effective but that requires a very careful administration), the list of events only refers to actions to be done in order to detect possible contraindications. Again, if CT scan shows an intracranial hemorrhages, that is one of the contraindications, "r-tPA Treatment" is switched off (from yellow to red) and less urgent events appear in the to-do list (see difference between lists (a) and (b) of Figure 1).

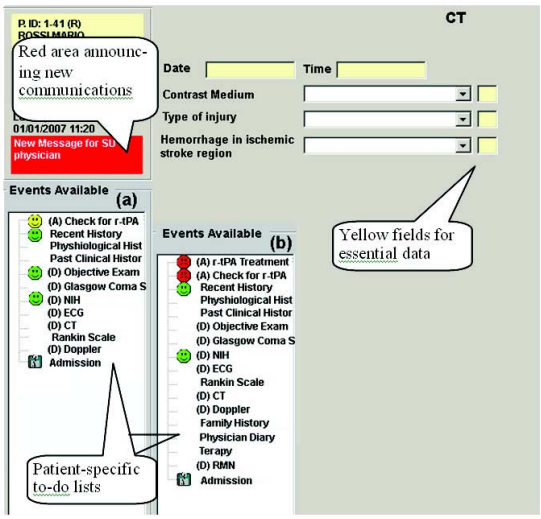


Figure 1 - A smart form of the Clinical Chart integrated with the CfMS. (a) and (b) are two different to-do lists for the same patient shown in two different points in time

Showing CfMS reminders

How to show GL reminders is a big issue in the implementation of a decision support system into a real healthcare setting. In general, pop-up windows are not well-accepted by physicians, because they are considered too intrusive. Thus, we preferred to rely upon the above described colored visualization of the to-do list: clicking the task icon, the data input form immediately appears. Moreover, a red area informs the users about new communications (red rectangle in Figure 1). This modality is minimally intrusive because users may open the communication box asynchronously, on their own will. The communication list, in chronological order, shows the priority level of each recommendation, its type (whether it is a diagnostic process, a treatment, a variable monitoring, or a message) and its positive or negative sign (if it is something to do or to rule out). Different users receive different communications according to their role (physician, nurse, physiotherapist, etc). Users may read communications for a specific patient or for all the patients currently admitted in the SU. If an answer is required, a checkbox appears near the text, allowing the receiver to say whether he/she agrees or not with the SPREAD recommendation. If a physician accepts a suggestion about a drug prescription, a record with the drug name, dosage and timing is stored and at the same time the CfMS takes care to inform the role "nurse", responsible for drug administration, sending a message with the useful details. Thus CfMS is also able to manage automatically a large amount of communication acts among personnel involved in patient care, making between-users interoperability more efficient by supporting medical and organizational knowledge sharing.

Ancillary tools

Natural language reporting

With the new system release, interpretation of GL rules is easier and safer but information is no more immediately available for nice printed free-text reports. To meet this physicians' need we developed a tool for the generation of such reports. Briefly, code descriptions are combined with associated notes (if any), according to grammatical rules, to generate the correct words and suitable introduction and conjunction sentences. The resulting reports are structured in distinct paragraphs for physiological history, recent and past clinical history, family-related history; information about past history is grouped by year and ranked in chronological order; once generated, the report may be edited by physicians before storing or printing.

Detecting and showing non-compliance

In a real-world setting, it is extremely difficult to detect non-compliance in real time. There are too many variables that could affect timely data storing and, consequently, timely automatic detection of non-compliance. This is particularly true in acute units. So, demanding that physicians justify their non-compliance has some drawbacks: first, detected non-compliance could be not real, i.e. an action has been performed but related data have not been entered in the system yet; second, users perceive the system not only a gentle reminder, but as a controller; eventually, if he's in a hurry, physician may have no time to write down justification. As a consequence, provided justification may be not reliable. For these reasons, in agreement with our clinical partners, we decided to manage non-compliance in a less invasive way. The point in time where we decided to accomplish this task is the patient discharge. This choice is based on the consideration that when physicians prepare patient's discharge, they are in team and they are not in a hurry. In addition, compiling the discharge letter requires summarizing the patient care process, and reasoning about it. Thus, presenting the non-compliance list at that moment has two advantages: overcoming the previous issues and facilitating bethinking about the case. In practice, at the patient discharge, a set of queries is activated on the patient record, and a report is produced. It shows the non-compliance list detected automatically, starting from the CCC, and gives the physicians the possibility to provide motivations that can be useful for further revision of the SPREAD GL (that is regularly updated every 2 years).

Results

The CfMS, after an intensive testing in our university laboratory, has been implemented in the SU (semi-intensive six beds unit) on April 2006. Three physician's and two nurse's workstations have been installed. Since the CfMS did not introduce dramatic changes in the end-user application, it has been accepted by healthcare personnel and entered the daily medical routine. We administered the Likert Scale SUS (Software Usability Scale [8]) to the five users of the Stroke Unit, obtaining an average value of 82.5/100, ranging from 80 to 85.

In order to evaluate the impact of the system in the daily activity, a set of queries has been developed on CfMS data, for checking both data completeness and GL compliance. Reports are produced by analyzing all medical data available for ischemic stroke patients admitted within a user-defined period. Reports contain statistics about:

- encoded data input: evaluation of the completeness of the minimum data set required by the GL;
- GL compliance: evaluation of the adherence to all the recommendations of the acute stroke phase;
- additional process variables, useful to capture the system utilization rate and modality, system intrusiveness and additional information about the clinical routine.

We compared the care processes of ischemic stroke patients admitted from April to November 2006 (113 patients managed with the help of the CfMS) with those of the 141 patients admitted in the same period of 2005, when the CCC was the same, but without decision support functionalities. Homogeneous periods avoid seasonal biases.

Table 1 - Encoded data input completeness

Event	April-November 2005	April-November 2006
Personal data	97%	98%
Physiological history	96%	97%
Recent history	50%	65%
Past history	94%	97%
Neuroradiology diagnosis	1%	39%
Neurosonologic exam	0%	0%
Cardiologic diagnosis	2%	35%
Diagnostic hypothesis	0%	35%
Objective examination	46%	56%
NIH stroke scale	76%	85%
Hematochemical exam	97%	97%
Rankin scale	35%	35%

Encoded data input

Completeness of encoded data input is increasing, since the CfMS introduction, for most of the events useful for GL interpretation (Table 1). The apparently astonishing result for neurosonologic examination highlights a double-input problem: this test is executed outside the SU and its result is automatically stored in the CCC as free text. This observation led to an update of the interface, which now shows, at the user's request, the free text, and prompts for encoding. Of course the best solution should be an update of the radiology department information system. In fact, the possibility of the CfMS of *identifying* and *quantifying* organizational problems is very important because it represents a sound starting point for further refinements of the hospital information system.

Guideline compliance

Of course completeness of encoded data affects the possibility of automatically evaluating compliance. Thus, due to poor completeness in the past, only few recommendations could be compared. In fact in many cases we have not enough encoded data to verify the recommendation and for other recommendations we could recognize too few eligible patients. Fortunately, sufficient data exist to compare compliance for three important recommendations: the SPREAD 10.5 (ASA in acute phase, grade A), 10.18 (Deep Venous Thrombosis prevention, grade B) and 10.7 (anticoagulant treatment with heparin, acute phase, grade D). Table 2 shows the number of eligible patients and the percentage of compliance for both periods, 2005 and 2006. For recommendation 10.7 the percentage of compliance remains the same and in both cases there is a total adherence to the SPREAD GL. There is a significant increase in compliance for recommendations 10.5 and 10.18.

Table 2 - GL compliance

SPREAD Recommendation	April-November 2005		April-November 2006	
	Eligible patients	% compliance	Eligible patients	% compliance
R 10.18	21	57%	21	86%
R 10.5	93	67%	48	71%
10.7	29	100%	8	100%

Additional process variables

Adoption of a WfMS has the main objective of improving efficiency and effectiveness of the organization processes; however, as a by-product, it also greatly facilitates obtaining interesting statistics on the system performance and on process variables. Since the system installation, 5643 records have been exchanged between the CCC and the WfMS components (average 36/day). These records concern to-do lists, messages, and clinical patients' data. Table 3 shows that the system is not too intrusive, since the amount of messages arriving to the communication box is very reasonable. The different percentage of therapy vs alert messages in the two patients groups (severe vs mild) is explained by the fact that severe patients have a longer hospital stay, therapy recommendations are concentrated in the first days, and then it is necessary to monitor side effects and complications. As another example, Table 4 shows, on the basis of encoded data entered in the CCC, the evaluation of the timing for the first two tasks that must be executed, as indicated by the SPREAD GLs, as soon as the patient is admitted: patient's history and diagnostic assessment (based on Rankin Scale, NIH Stroke Scale, Neuroradiologic and Neurosonologic examination). Although computed times to accomplish tasks are obviously biased toward high values (data input could be delayed), these data highlight and quantify possible problems in performing some diagnostic tests. It's worth noting that workflow process data allow to detect bottlenecks due to lack/misuse of resources (both technological and

Table 3 - Message exchange figures

	Average hospital stay	n. of messages /patient	n. of messages /day/patient	Therapy suggestion messages	Alert messages
Severe patients (MRS[§]>2)	10.55	11	1.062	14%	86%
Mild patients (MRS<=2)	6.61	6	0.953	30%	70%

§ Modified Rankin Scale

Table 4 - Evaluation of the clinical routine timing

	Elapsed time between admission and complete history		Elapsed time between admission and diagnostic assessment	
	<6 hours	<24 hours	<6 hours	<24 hours
Severe patients	67%	85%	15%	50%
Mild patients	71%	94%	33%	56%

human), thus allowing to discover the causes of the above cited problems.

Conclusions and future developments

Despite the big potential of Information and Communication Technology to improve healthcare organizations' outcomes (in terms of efficiency and/or effectiveness), there are still several problems in achieving successful real-world implementations. It is estimated that 60-70% of all software projects fail [9]. In this paper we have described the implementation of a decision support system aimed at improving compliance to GLs for stroke management. Results of the first eight months of the system use in the Stroke Unit are very encouraging. We think that key factors to the system success were manifold. First of all, users' commitment was high: the need of improving clinical practice was manifested by the neurological medical community, and all the system functionalities have been designed together with physicians. Second, the principle of minimum intrusiveness has been followed, and we took care in introducing minimal changes to the interface that users were familiar with. Last, we strongly collaborated with third partners providing the CCC, in order to achieve full system integration, and promptly intervene on users' feedback about system bugs and flaws. About future developments, we will focus on advanced technologies for data input. As a matter of fact, a well-known bottleneck for the optimal use of decision support systems is the difficulty to insert data in real time. We mentioned that some statistics could be biased by the delayed interaction. Ubiquitous computing technologies, provided for example by PDA, vocal interaction, and wireless connections, could greatly increase the CfMS performance and results reliability.

Acknowledgements

This work has been partially funded by projects FIRB (Italian Ministry of University and Research) and STAGE (Italian Ministry of Health).

References

[1] Kohn LT, Corrigan JM, and Donaldson MS, eds., To Err Is Human: Building a Safer Health System (Washington: National Academy Press, 1999).

[2] Micieli G, Cavallini A, Quaglini S; for the GLADIS Study Group. Guideline compliance improves stroke outcome: a preliminary study in 4 districts in the Italian region of Lombardia. Stroke. 2002,33(5):1341-7.

[3] Peleg M, Tu S, Bury J, Ciccarese P, Fox J, Greenes RA, Hall R, Johnson PD, Jones N, Kumar A, Miksch S, Quaglini S, Seyfang A, Shortliffe EH, Stefanelli M. Comparing computer-interpretable guideline models: a case-study approach. J Am Med Inform Assoc. 2003 Jan-Feb;10(1):52-68.

[4] 2006 Workflow Handbook Layna Fischer Ed., Future Strategies Inc.Lighthouse Point, FL, USA, 2006.

[5] SPREAD 2005 - Stroke Prevention and Educational Awareness Diffusion (2005) Edition of march 15th, www.spread.it.

[6] Baud RH, Weber P, Lovis Ch, Coding in context. PCSE 2001 Brugge.

[7] Panzarasa S, Quaglini S, Micieli G, Pernice C, Stefanelli M, Pessina M. Workflow technology to enrich a computerized clinical chart with decision support facilities. Proc. AMIA 2006, November 10-15, Washington, 2006.

[8] Brooke J. (1996) SUS: a "quick and dirty" usability scale. In P W Jordan, B Thomas, B A Weerdmeester & A L McClelland (eds.) Usability Evaluation in Industry. London: Taylor and Francis.

[9] Ammenwerth E, Iller C, Mahler C. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. BMC Medical Informatics and Decision Making 2006, 6:3.

Address for correspondence

Silvia Panzarasa
Consorzio di Bioingegneria e Informatica Medica
UPIT,
Via Ferrata,
1 - 27100 Pavia (Italy).

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Towards a Decision Support System for Optimising Clinical Pathways of Elderly Patients in an Emergency Department

Marc Cuggia [a], Delphine Rossille [a], Aude Arnault [b], Jacques Bouget[c], Pierre Le Beux [a]

[a] *EA3888, Faculté de Médecine, Université de Rennes1, France,* [b] *ENSAI, Bruz, France*
[c]*Service de l'accueil des Urgences, CHU Pontchaillou, France*

Abstract

Data stored in Healthcare Information Systems correspond potentially to a vast source of information for supporting decisions in management or public health issues. The presented study illustrates clinical data valuation, through the analysis of clinical pathways of elderly patients at the Emergency Department (ED) of Rennes hospital. Method: Relevant data were extracted from the Emergency Department database. Several analysis (e.g., cusum method) and representation tools (e.g., Graphviz) were used to study the patients' pathways, the dynamics of flows and the patients' characteristics. Results: 4951 admissions were analysed and visualized. The representations provided a synthetic, global and comprehensive view of the department activities, to the satisfaction of the clinicians. Limitations of the ICD-10 coding of the diagnoses at the ED were pointed out as well as syntax and semantic interoperability issues. A solution is proposed for automating and scaling the Decision Support System.

Keywords:

management information systems; emergency service, hospital; critical pathways; decision making, computer-assisted

Introduction

The first departments to introduce and use computers in French public hospitals were administration and finance departments. The main purpose of collecting these medico-economic data was to help monitor the hospitals. Specific analytical systems were developed, some using for instance data mining methods, and this allowed data to be used as evidence in annual reports to determine budget allocation. One of our hospitals' main challenges is now to computerize clinical departments, hence healthcare processes. Thanks to the development of healthcare tools (i.e., for medical prescription, medical records, nursing records), it is possible to collect clinical data in the Hospital Information System's (HIS) databases on a day-to-day basis. These data are exhaustive, up-to-date and in a storage format facilitating their exploitation. They potentially represent a vast source of information and knowledge of prime interest for Public Health studies within hospitals. From the HIS databases, analysis and decision support tools could be developed to address Public Health issues: quality of care, hospital organisation and epidemiological research studies.

The objective of this paper is to illustrate the exploitability of the clinical and medico-economic data produced by the HIS for decision support in hospital management by studying the activity of a specific department. Several problems were identified.

The present study focuses on the care of elderly patients (aged 75 or over) in the emergency department (ED) of Rennes hospital. It is well known that the current conditions of hospital care of elderly patients are not satisfactory: waiting time at the ED is usually too long and patients are not always transferred to the hospital departments relevant to their conditions. The situation becomes worse in times of crisis, such as during the long spell of very hot weather in the summer of 2003 when the ED could not handle the patients flow and an unexpectedly high mortality rate was reached. In this context, the Emergency Department is in demand of a better understanding of its activity related to elderly patients, especially in terms of patient pathways.

This paper presents the results of the study from the decision support point of view. The aim is to provide feedback to the Emergency Department in the form of comprehensive reports of quality and effectiveness of care. In addition, the study should contribute to a better understanding of the delivery of care to elderly patients within the hospital in a public-health perspective.

The paper will focus on three areas of interest for the ED: the characterisation of patients, dynamics of flows and patients' pathways within the Emergency Department (that is the patients' flow within the ED). The three points will be discussed from a public health perspective, stressing the lack of formal semantics in the medical concept representations in use at the hospital.

Background

Several papers describe differentways to exploit data produced by ED's information systems for decision support in medical, management and public-health issues.

While knowledge-based systems aim at supporting medical decisions in the ED at the time of care (e.g.,. QMR, Iliad), other focus on management issues, such as the coordination, resource allocation and documentation aspects of

the Emergency Department operations [1]. Some specialize in the triage process, helping to prioritize incoming patients. For instance, the MET-AT system [2] provides triage plans for acute paediatric abdominal pain. Others propose methods (based on the cumulative sum chart) to manage hospital bed occupancy crisis [3].

Other papers relate experiences of exploiting data for public-health purposes. In Sydney [4], Australia, during the 2003 Rugby World Cup in a context of bioterrorism threats and emergent diseases, national public health authorities tested a near real-time syndromic surveillance system. In this system, EDs in the Sydney metropolitan area automatically transmitted surveillance data from their existing information systems to a central database in near real-time. This information included patient demographic details, presenting problem and nursing assessment entered as free-text at triage time, physician-assigned provisional diagnosis codes, and status at departure from the ED. Automated processes were used to analyse both diagnosis and free text-based syndrome data and to produce web-based statistical summaries for daily review. An adjusted cumulative sum (cusum) and Bayesian methods were used to assess the statistical significance of trends. In France, at a national level, the Heat Health Watch Warning System [5] was developed in 2003 to anticipate heat waves that may result in a large excess of mortality. The system was developed on the basis of a mortality analysis, healthcare activities in EDs and meteorological data. Several meteorological indicators were tested in relation to levels of excess mortality. The system requires close cooperation between the French Weather Bureau (Meteo France), the National Institute of Health Surveillance (InVS) and the Ministry of Health. The system is supported by a panel of preventive actions, to prevent the sanitary impact of heat waves.

Such systems illustrate the current interest in exploiting healthcare data for public health and management purposes. They rely on information whose format and sources are heterogeneous. The collection of clinical data from EDs is meant to improve the delivery of care for patients in primary care.

Materials and methods

The choice of the emergency department was motivated by its central dispatching role in the hospital and the daily contribution of medical records in its clinical database. Elderly patients only were selected as they represent an ever increasing part of the population in developed countries. Most of them have poly-pathologies.

The emergency department is structured in three medical wards: medicine, surgery and small traumatology. A proximity unit (named Duhamel ward) is specific to the ED. Patients staying in the Duhamel ward (theoretically for a maximum of 24 hours) either need to stay in observation or wait for an available bed in an appropriate hospital department.

Materials: Data were extracted from the ED clinical database (RESURGENCE) and the medico-economic database (DRG) produced by the HIS. The clinical medical records include general data (e.g., sex, age), data on arrival, departure and transfer (e.g., arrival mode, arrival data), data on the follow-up at ED, on the diagnoses and medical status. Medico-economic data include ICD-10-coded diagnoses, medical procedures (coded with a French classification named CDAM), date of hospitalisation and the various hospitalisation departments where the patient went. Hence only patients transferred in a hospitalization department appear in the medico-economic database. All the extracted data are limited vocabularies. Only data of elderly patients were selected. Over a period of 6 months (from October 2005 to may 2006), 32883 patients came to the emergency department, 4951 were aged 75 or over.

Method: After extracting all the data from the HIS databases (SQL requests), relevant data to the study's objectives were identified and their quality control checked through statistical analysis (SAS). The analysis methods and representational views were chosen to reach the objective of providing useful, easy-to-read and synthetic views of the ED activities.

Descriptive statistics provided characterisation of the population, more specifically on the length of stay at ED and the diagnoses. To study the dynamics of flows, the cusum method was used on the admission rates. A cusum chart [3] is the cumulative sum of consecutive differences between an individual measurement and a selected target. It reveals trends otherwise not obviously detectable in the corresponding time-series data. This analysis technique has been applied to manage hospital bed [3] or to monitor trends in hospital-acquired infection. For visualizing patient pathways within ED, the data were pre processed according to a new arrangement of the operational structures. Patterns of patient pathways were displayed using the visualization tool Graphviz (http://www.research.att.com/sw/tools/graphviz/). The chosen visualization tool was aimed at providing a global representation of the most frequent (and typical) pathways for the population, giving insights on the use of the different ED wards as well as on the waiting times.

Results

The demographic study shows that the admitted elderly population is representative of the elderly in France and corresponds to 17.9% of the total number of patients admitted at the ED, slightly higher than national percentages (15 to 17%).

Characterisation of the population

The **ED length of stay** (expressed in hours) depends strongly on the week day of arrival and the patient's medical status at arrival. The medical status is expressed in a scale from 1 (stable prognosis) to 5 (vital prognosis in danger, immediate intensive care required). The highest mean length of stay is found on Sunday (Figure 1), a week-end day during which the medical manpower is currently reduced. Not surprisingly, the length of stay is correlated with the patient's medical status ($p>0.005$), to the exception of cases of extremely bad medical status corresponding to patients with potentially lethal conditions (Figure 2).

From the distribution of the **diagnoses** as a function of the ICD-10 main chapters (Figure 3), the three most repre-

sented principal diagnoses at the ED are Chapter XVIII "Symptoms,

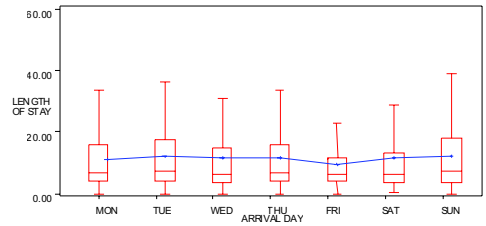


Figure 1 - Tukey box for the ED length of Stay vs. the week of the day at arrival

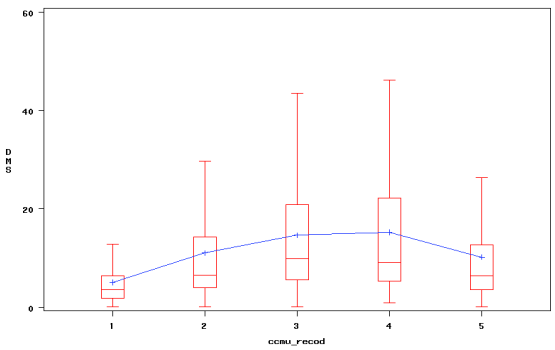


Figure 2 - Tukey box for the ED length of Stay vs. the patient's medical status at arrival

signs and abnormal clinical and laboratory findings, not elsewhere classified" (22.7%), Chapter XIX "Injury, poisoning and certain other consequences of external causes" (20.9%) and Chapter IX "Diseases of the circulatory system" (13.3%).

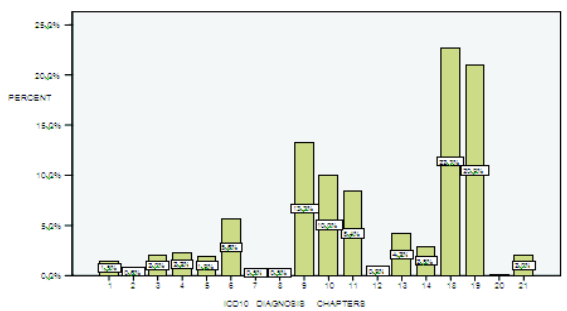


Figure 3 - Distribution of the principal diagnoses at ED

Table 1 details diagnoses for the three principal diagnoses at ED. It shows for instance that 48.5% of the diagnoses related to Chapter XVIII correspond to "General symptoms and signs".

Dynamics of flows

Figure 4 represents the number of admissions at ED with respect to the date, for the considered period of time. The time-series data appear to be very complex to interpret. Using the cusum method provides a comprehensive and global view of the variations (Figure 5). Indeed, the slope changes correspond to admission rate higher than the mean (that is the selected target). Hence three periods show an increased admission rate: in December, February and April.

Chapter 18		Chapter 19		Chapter 9	
48.5 %	General symptoms and signs	26.4 %	Injuries to the hip and thigh	41.8 %	Cerebrovascular diseases
14.5 %	Symptoms and signs involving the circulatory and respiratory systems	20.1 %	Injuries to the head	29.9 %	Other forms of heart disease
11.3 %	Symptoms and signs involving nervous and musculoskeletal systems	10.4 %	Injuries to the shoulder and forearm	11.8 %	Diseases of arteries, arterioles and capillaries

Table 1 - Detailed diagnoses for the three principal ICD-10 diagnoses at ED

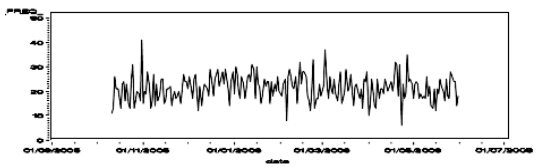


Figure 4 - Number of admissions per date

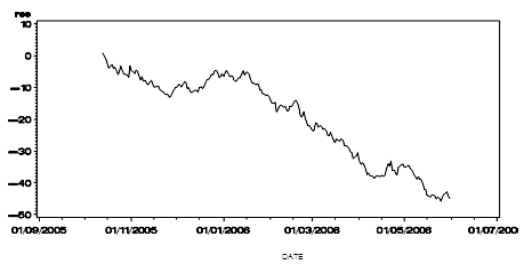


Figure 5 - Cusum chart of the number of admissions per date

Patient pathways

Figure 6 provides a visual representation of the patient pathways within the Emergency Department. Only pathways of type <admission – step 1 - step 2> are considered. They represent alone 82.7% of the total patient pathways.. A "step" consists of one of the four ED services or an exit mode. The pathways not represented correspond to complex pathways of more than three steps, or to pathways taken by very few patients. At each step is given the median waiting time. On each branch of the graph the percentage of the patients following this path between the two nodes is displayed, 100% referring to the father node. The most followed pathway corresponds to <admission – medicine service – exit>(49%, represented in red). Then come the pathway <admission – surgery service – exit> (30%, in grey), far behind <admission – traumatology

service – exit> (2.42%, in yellow) and <admission – exit> (0.67%, in blue).

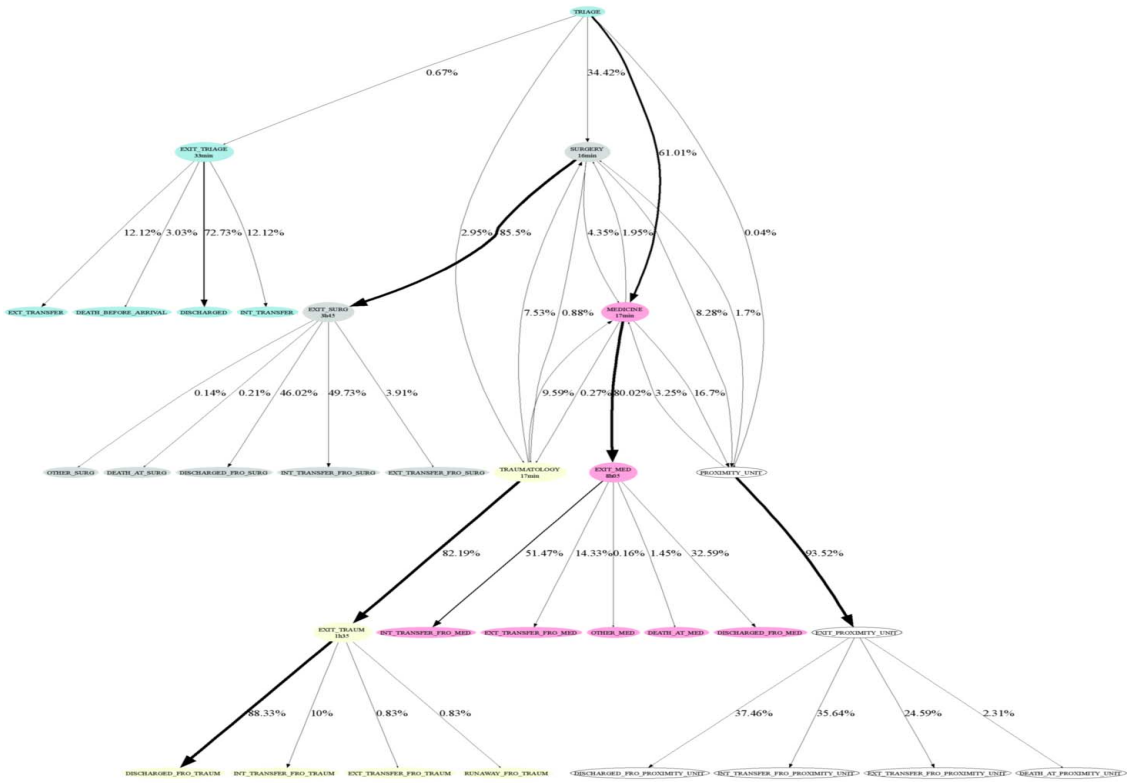


Figure 6 - Patients' pathways at the Emergency Department

Discussion

The study showed that the vast source of medical data stored for healthcare purposes in the HIS databases can provide valuable insights to the clinicians in a comprehensive and straightforward manner. Clinicians at the Emergency Department were very satisfied with this study of elderly patients and with the chosen representation tools. No actual satisfactory survey was done, but the team responsible for the optimisation of the patients' flows within the hospital has decided to use this work as it provides a quantitative measure of the ED activities, especially in terms of patients' pathways.

Thanks to these representation tools used, the patients' characteristics, the dynamics of flows and the patients' pathways provide a multifaceted and complete view of the Emergency Department easily readable by the clinicians for decision support purposes, as for improving the department structural organisation and service management.

Dynamics of flows : The temporal analysis (cusum method) provides a yearly view of the dynamics of flows. It is complemented with the periodic representation per week days (Tukey boxes). This has led the Emergency Department to look at increasing medical manpower during week-ends, in order to face the higher patients flow.

Patients' pathways : Graphviz allowed a relatively clear representation of the patient pathways. However many pathways had a frequency of almost one. It was decided to represent graphically only the most important pathways in terms of frequency. Hence a subjective threshold was applied to patients' pathways. The Galois lattice calculates the threshold based on statistics [6]. However the lattice structure and its graphical representation are much more difficult to interpret than the Graphviz representation that was proposed in this paper.

Medical characterisation of the population : Limitations of the exploitation of healthcare data appear to lie in the use of ICD-10 coding for diagnoses. As shown in Table 1, the most frequently-used code for diagnoses at the ED corresponds to "Symptoms and general signs" of Chapter XVIII of ICD10 "Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified". Indeed determining precise diagnosis at the ED is often difficult as for instance several complementary exams may be required during hospitalisation. Priority is first of all to deal with acute situation, to start symptomatic treatment and to transfer the patient to the more appropriate hospitalisation department when needed. It is not surprising that ED clinicians often code the symptoms (s.a. abdominal pain, migraines) instead of precise diagnoses. However the ICD-10 is primarily intended to classify diseases while the

symptoms are described only in one chapter out of 22, therefore not being linked in any way (into the ICD-10 structure) to the underlying disease.

To solve this problem of coherency of ICD-10, a recent work suggests SNOMED CT to be the coding reference. Indeed, a comparative study within the UMLS shows that SNOMED CT covers better the needs for coding medical records in the Emergency Department [7]. Another paper [8] shows that better semantic coherency is possible with SNOMED CT when comparing diagnoses codes at the different steps of the patient pathway in the hospital.

Perspective

To be scalable to the hospital level, regional or even national level, such study requires the setting up of an automatic Decision Support System. With the presented preliminary work, several issues were brought to light on the system's architecture and the syntax and semantic interoperability within the HIS.

First a data warehouse should be built in order to gather all the heterogeneous data extracted from the HIS. This approach facilitates the transformation and connexion of healthcare data for decision support purposes. Calculating decisional indicators can be very time and resource consuming and having already all the required data in the proper format in the data warehouse will guaranty the complete availability of healthcare resources for medical activities. The process of loading data from heterogeneous sources into the data warehouse assumes syntax interoperability issues to be solved. This can be achieved by using international exchange standards, such as the HL7 [9] and should be easier if the HIS is built on an SOA architecture (using web services)[10].

However semantic interoperability issues remain to be considered. For instance, taking the case of a Decision Support System at the hospital level, the adequacy between the diagnoses coded at the ED (often "symptoms") and the diagnoses coded at the hospitalisation departments is not straightforward as shown in the study. This could be solved in two steps. First the "symptom" would be related to an appropriate "disease" as hospitalisation departments are specialised to cure specific pathologies (s.a. cardiology or pneumology departments). The SNOMED CT, with its semantic relationships between symptoms and diseases, could be used [8]. The second step consists of checking the adequacy between the hospitalisation department and the patient's disease. This task is not straightforward and several ambiguities exist. For instance, the internal medicine department is in charge of a large number of different pathologies. Overlapping of competency between two separate departments may also occur. For instance, a patient with an acute oedema of the lung may be directed either to the cardiology or the pneumology department. One solution is to describe automatically the competency fields of departments based on their economic activities, for instance by using the main Diagnosis Related Groups (DRGs) of each department.

Conclusion

The objective of the paper is to show, through the study of the delivery of care to elderly patients at the Emergency Department, how data produced by healthcare processes are valuable for decision support systems. This approach is related to the data mining approach whose objective is to extract relevant knowledge for the hospital management.

The results indicate that providing a synthetic and global view of the department's activities, thanks to appropriate representation tools comprehensive to clinicians, is particularly relevant for the department's managers and clinicians.

Finally, it is explained that for scaling the application to the hospital level, issues of syntax and semantic interoperability have to be considered. A solution is proposed based on SNOMED CT for its semantic medical representation and DRGs to allow the automatic description of the competency fields of clinical departments.

References

[1] R. Ceglowski, L. Churilov, J. Wassertheil, "Facilitating decision support in Hospital Emergency Departments: A process-oriented perspective", in *Proceedings of the 13th European Conference on Information Systems*, Regensburg, Germany, 2005

[2] K.J. Farion, W. Michalowski, S. Rubin, S.Wilk, R. Correll and I. Gaboury, "Prospective evaluation of the MET-AP system providing triage plans for acute pediatric abdominal pain", *Int J Med Inform*., 2007 Feb 21.

[3] C. M. Burns, C. J. Bennett, C. T. Myers, and M. Ward, "The use of cusum analysis in the early detection and management of hospital bed occupancy crises," *Med J Aust*, vol. 183, pp. 291-4, 2005.

[4] D. J. Muscatello, T. Churches, J. Kaldor, W. Zheng, C. Chiu, P. Correll, and L. Jorm, "An automated, broad-based, near real-time public health surveillance system using presentations to hospital Emergency Departments in New South Wales, Australia," *BMC Public Health*, vol. 5, pp. 141, 2005.

[5] M. Pascal, K. Laaidi, M. Ledrans, E. Baffert, C. Caserio-Schonemann, A. Le Tertre, J. Manach, S. Medina, J. Rudant, and P. Empereur-Bissonnet, "France's heat health watch warning system," *Int J Biometeorol*, vol. 50, pp. 144-53, 2006.

[6] N. Jay, A. Napoli, and F. Kohler, "Cancer patient flows discovery in DRG databases," *Stud Health Technol Inform*, vol. 124, pp. 725-30, 2006.

[7] J. C. McClay and J. Campbell, "Improved coding of the primary reason for visit to the emergency department using SNOMED," *Proc AMIA Symp*, pp. 499-503, 2002.

[8] P. L. Elkin, A. P. Ruggieri, S. H. Brown, J. Buntrock, B. A. Bauer, D. Wahner-Roedler, S. C. Litin, J. Beinborn, K. R. Bailey, and L. Bergstrom, "A randomized controlled trial of the accuracy of clinical record retrieval using SNOMED-RT as compared with ICD9-CM," *Proc AMIA Symp*, pp. 159-63, 2001.

[9] K. M. Coonan, "Medical informatics standards applicable to emergency department information systems: making sense of the jumble," *Acad Emerg Med*, vol. 11, pp. 1198-205, 2004.

[10] J. Mykkanen, A. Riekkinen, M. Sormunen, H. Karhunen, and P. Laitinen, "Designing web services in health information systems: From process to application level," *Int J Med Inform*, 2006.

Address for correspondence

Dr Marc CUGGIA
Département d'information médicale - CHU Pontchaillou
2, rue H. Le Guilloux - 35033 Rennes Cedex

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Ontology-based Modeling of Clinical Practice Guidelines: A Clinical Decision Support System for Breast Cancer Follow-up Interventions at Primary Care Settings

Samina R. Abidi, Syed SR. Abidi, Sajjad Hussain, Mike Shepherd

NICHE Research Group, Faculty of Computer Science, Dalhousie University, Halifax, Canada

Abstract

Breast cancer follow-up care can be provided by family physicians after specialists complete the primary treatment. Cancer Care Nova Scotia has developed a breast cancer follow-up Clinical Practice Guideline (CPG) targeting family physicians. In this paper we present a project to computerize and deploy the said CPG in a Breast Cancer Follow-up Decision Support System (BCF-DSS) for use by family physicians in a primary care setting. We present a semantic web approach to model the CPG knowledge and employ a logic-based proof engine to execute the CPG in order to infer patient-specific recommendations. We present the three stages of the development of BCF-DSS—i.e., (a) Computerization of the paper-based CPG for Breast Cancer follow-up; (b) Development of three ontologies—i.e., the Breast Cancer Ontology, the CPG ontology based on the Guideline Element Model (GEM) and a Patient Ontology; and (c) Execution of the Breast Cancer follow-up CPG through a logic-based CPG execution engine.

Keywords:

clinical practice guidelines, breast cancer,
decision support system, medical ontology, semantic web

Introduction

Clinical Practice Guidelines (CPG) entail medical knowledge intended for clinical decision-making and standardization of clinical practice [1, 2]. Despite the potential benefits of CPG, reviews show that CPG are underutilized in clinical practice [3, 4], largely due to problems associated with their dissemination to physicians [2, 5]. CPG computerization involves the modeling and conversion of a paper-based CPG into an electronic and executable format that can both be accessed by physicians and be embedded within clinical decision-support systems at the point of care. CPG guided decision support systems are particularly useful in clinical settings where non-specialist health practitioners, such as family physicians or nurses, are required to deal with complex or unusual cases. In such situations, CPG based decision support systems can guide the healthcare practitioner's actions and suggest appropriate recommendations. One such situation is the discharge of Breast Cancer (BC) follow-up care by family physicians. Note that in Nova Scotia follow-up care sis currently being provided by cancer care specialists at tertiary care centers.

Recent advancements in BC treatment have significantly improved the rate of BC survivors in Nova Scotia. The follow-up care for patients in remission entails periodic visits for history, physical exams and mammogram surveillance [6]. Although specialized cancer clinics provide long-term follow-up care, there is a case for formally involving family physicians in breast cancer follow-up care. In fact, trials conducted in Canada and Britain show that family physician offer a viable alternative to specialized care clinics for offering follow-up care to women who are in remission from breast cancer [6]. However, for most family physicians BC follow-up care is a new and added responsibility, therefore they need clear clinical guidelines to effectively perform the follow-up activities, make correct decisions and provide the right recommendations. The Canadian Steering Committee on CPG for the Care and Treatment of Breast Cancer has developed and recently updated the guideline on follow-up care after treatment for BC [7] with special emphasis on the needs of primary care physicians. The challenge was to disseminate the CPG to the family physician and to integrate it within his/her clinical workflow so that the CPG is seamlessly executed whenever a patient undergoes BC follow-up in a primary care setting.

In this project, we collaborated with Cancer Care Nova Scotia to address the abovementioned challenges by promoting the (knowledge) translation of the CPG for the Care and Treatment of BC to family physicians, to support them in the delivery of BC follow-up care and patient education at their clinics. This will reduce the workload of specialist cancer centers within Nova Scotia. Our approach was to develop an interactive Decision Support System (DSS) that enables family physicians to (a) access and utilize the said CPG at the point of care to provide standardized follow-up care; and (b) offer customized patient educational information targeting disease management, lifestyle behaviours and psychosocial support.

In this paper, we present an ontology-based Breast Cancer Follow-up Decision Support System (BCF-DSS) based on the CPG for the Care and Treatment of Breast Cancer. We

take a semantic web approach to model the CPG knowledge and to reason over the ontology to provide 'trusted' CPG-driven recommendations. We have developed three ontologies: (a) *CPG ontology* that models the structure of the CPG based on the Guideline Element Model (GEM); (b) *Breast Cancer Ontology* that represents the medical knowledge encapsulated within the CPG and general BC related concepts; and (c*) Patient Ontology* that models the patient's parameters. The ontologies are developed using Protégé and are in OWL format. We have developed a logic-based reasoning engine that reasons over the knowledge from these three ontologies. Our BCF-DSS allows family physicians to collect patient data and assists them to make CPG mediated decisions, recommendations and referrals for BC survivors. We present the three stages of the development of BCF-DSS—i.e. (a) Computerization of the paper-based CPG for the Care and Treatment of Breast Cancer; (b) Development of the ontologies, in particular the Breast Cancer Ontology; and (c) Execution of the BC follow-up CPG through a logic-based CPG execution engine.

Computerization of BC follow-up CPG

Computerization of the CPG involved (a) selection of a CPG modeling formalism; and (b) capturing and representing the CPG knowledge based on the modeling formalism. We selected the Guideline Representation Model (GEM) to model the BC follow-up CPG. GEM is based on XML that renders it operable in a semantic web environment by allowing semantically salient indexing and searching of CPG knowledge. We used the GEM Cutter tool to annotate the BC follow-up CPG with GEM tags (or elements). The conversion task involved determining the function of a specific CPG text and annotating it using the relevant GEM tag. It may be noted that GEM constitutes 100 tags covering a wide variety of concepts. For our purposes, the most salient concepts were the 'Knowledge Components' that store and categorize the knowledge present in a CPG. The knowledge components have subcomponents called recommendations which are categorized as either *imperative*—i.e., directed towards entire target population, or *conditional*—i.e., act on the decision variables and results in appropriate actions.

From a knowledge modeling perspective the main challenge was to resolve the (medical and semantic) ambiguities inherent within the BC follow-up CPG. To resolve the ambiguities we (a) consulted with BC oncologists, in particular the author of the BC follow-up CPG; (b) reviewed available literature; and (c) applied our personal clinical experience. Examples of ambiguities included phrases such as, 'vaginal bleeding is present in the absence of obvious cause', 'physiological causes of fatigue' and 'other risk factors of osteoporosis'. The phrase 'vaginal bleeding in the absence of obvious cause' was resolved to the term 'Menstruation', and other ambiguous phrases were similarly resolved by mapping them to explicit concepts. Finally, through the use of GEM we managed to create an executable representation for the BC follow-up CPG.

Development of breast cancer ontology

The BC ontology models the knowledge encapsulated within the BC follow-up CPG. We used Protégé ontology editing environment to build our BC ontology in OWL (Web Ontology Language) using Protégé OWL.

The BC ontology is largely derived from the contents of the knowledge components—i.e. the 'Imperative' or Conditional recommendations—in the GEM representation of the BC follow-up CPG. More specifically, the conditional recommendation element, which comprises sub-elements such as 'decision.variable', 'action' and 'logic' elements, was used to develop the BC ontology. Given conditional recommendations, the challenge was to identify the decision variables, the actions to be taken and the Boolean logical operations in the recommendations, so that the resultant ontology was compatible with our logical reasoning engine. In this regard, two design constraints were addressed: (a) Our CPG execution engine does not processes statements containing 'OR' and 'NOT' logical operators. Therefore a rule such as:

> "**IF** *age* >65 **OR** *family history of* osteoporosis **OR** *menstrual status of* premature menopause due to treatment,
>
> **THEN** screen with bone mineral density and treat accordingly with bisphosphonates",

was required to be decomposed into three smaller rules, each with a single decision variable so that the OR operator was eliminated; and (b) The BC ontology classes that have multiple domains or ranges could not be executed safely. Therefore, we ensured that all properties have a single domain and range.

Specifying BC ontology classes

Considering the above constraints we defined eight main classes, namely; Patient_Type, Physician_Type, Illness, Menstrual_Status, Recommendation, Symptom, Diagnostic_Tests, Treatment, Age, Risk_Factor, Weight_Status and Patient_Wish Next, we specified the disjoint classes, where classes are disjoint when an individual cannot be an instance of more than one of these classes. In the BC ontology the only classes which are not disjoint are the Recommendation and Diagnostic Tests, and Recommendation and Treatment since they share some instances, for example "Screening with bone mineral analysis" is an instance of two classes i.e., Diagnostic_Test and Recommendation; and "Bisphosphonates" belong to class Treatment and Recommendation.

Specifying properties for the BC ontology classes

Properties for patient class

The *'Patient_Type'* is the most important class because most conditional recommendations are targeted towards the patient. To specify different patient types we defined a range of properties. The patient properties represent the patient profile—i.e., an instance of the class *Patient_Type*. We defined object properties to establish link between the classes so that recommendations can be associated with a patient profile. The class *Patient_Type* has the most object

properties with their domain being *Patient_Type* but their range includes instances from other classes in the BC ontology, for example, the properties 'has_*history_of'* and *'has_illness'* have individuals of the class Illness as their range.

In our CPG execution engine most of the properties are treated as *decision variables* that serve as the premise of a logical rule. The conclusion of the rule is an *action variable* that corresponds to a recommendation, treatment, or statement directed towards a patient. To account for the patient-centric action variables, we specified two *Patient_Type* properties—i.e. *is_Recommended* and *possible_cause_can_be* with an unspecified range as their range can be any individual from any class. The action variable *is_Recommended* refers to the any recommendation, diagnostic test or treatment suggested to a patient. The action variable *possible_cause_can_be* provides the physician information regarding the cause of certain sign or symptom, for example the CPG statement "emotional distress, may be the underlying cause of subjective complaints of impaired cognitive functioning."

Properties for other classes

The properties for the other classes were derived from conditional statements in the recommendation element of the GEM representation of the BC follow-up CPG. These conditional statements specify variables that non-patient specific data, for example consider the CPG statement "If the purpose is to detect distant metastasis, then routine lab and radiographic exam should not be carried out". Such as statement was modeled by the class 'Diagnostic_Test' through two properties; *has_purpose_to_detect* whose range is Illness and *test_apply_to* whose no specified range. In total, we specified 40 properties for all other classes excluding the patient class.

Properties for statements having the not logical operator

Modeling of some statements in the BC follow-up CPG required the use of the 'NOT' operator. For instance consider the statement, "When such bleeding (vaginal bleeding) is present in the absence of obvious cause, endometrial biopsy should be carried out". Here the phrase 'in the absence of obvious cause', really means *not* obvious cause. We handled such situations by specifying a new property, for instance *is_not_caused_by,* for the class Symptom. Note that the rationale for creating such as property is because our execution engine is unable to handle the 'NOT' logical operator.

Specifying property characteristics

Certain properties such as *has_age*, *has_weight_status* and *has_menstural_status* are functional properties since a patient can have only one age, weight status (i.e., can either be over-weight, under-weight or have correct weight) and menstrual status (i.e., can either be premenopausal, postmenopausal or premature menopause due to treatment). Most of the properties are not functional and allow multiple values. We also specified some inverse object properties such as 'is_*recommended_for_illness*' which is the property of class 'Treatment' is the inverse property of '*is_treated_by*' which is the property of class 'Illness'. This means that 'Bisphosphonates' *is_recommended_for_illness* 'Osteoporosis' and 'Osteoporosis' *is_treated_by* Bisphosphonate. Note that Osteoporosis is the individual of the class Illness and Bisphosphonate is the individual of the class Treatment.

Specifying individuals (instances)

In the next step of BC ontology development we specified the individuals (or instances) from the conditional recommendations of the BC follow-up CPG. For example, individuals for class Symptom include Anxiety, Back_Pain, Cognitive_Impairment, Fatigue, Impaired_sexual_function, Menopausal_Symptoms, Vaginal_Bleeding, and Vaginal_Dryness. The class Patient_Type has the most individuals since each recommendation is valid for a patient with a particular set of clinical characteristics, thus each patient type refer to a particular patient profile in accordance to the said CPG.

Specifying relationships between the classes

The relationships among different classes were modeled using the class properties, and can be best understood by the following example. In order to model the recommendation statement "When such bleeding (vaginal bleeding) is present in the absence of obvious cause endometrial biopsy should be carried out". Here, "obvious cause of bleeding" means Menstruation. We established a relationship between the classes *Patient Type, Symptom, Menstural_Status* and *Diagnostic_Test* as follows: Patient_Type_12 is an individual of class Patient Type who has Vaginal_Bleeding which is the value for its object type property *has symptom*. Vaginal_Bleeding is also an individual of class Symptom, which has an object type property called *is_not_caused_by*, whose value is Mensturation_or_Obvious_cause, which in turn is an individual of another class called Mentrual_Status. The class Menstrual_Status has an object property called *ms_apply_to_diagnostic_test* whose values in this case will be Endometrial_Biopsy which is an instance to the class Diagnostic_Test (See Figure 1). By adding the property *is-not-caused_by* we ensured that the recommendation is logically valid i.e. endometrial biopsy is not recommended whenever the patient has vaginal bleeding. In this way we are able to establish an inter-class relationship that can be used to infer that if a patient has vaginal bleeding and bleeding is not caused by menstruation or any other obvious cause, then endometrial biopsy is the recommended test.

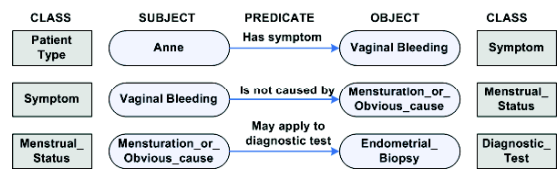


Figure 1 - RDF Triples depicting the relationships between classes Patient_Type, Symptom, Menstrual_Status, and Diagnostic_Test to model a recommendation

CPG Execution Engine

In keeping with our Semantic Web approach we developed a CPG Execution Engine (CPG-EE) that constitutes (a) multiple ontologies to model the domain and CPG knowledge; (b) a logic-based proof engine that leverages the ontologies and CPG specific rules to infer CPG mediated recommendations; and (c) a justification trace to describe the rationale for the inferred recommendations; this is to establish 'trust' in the proposed recommendations (Figure 2). The CPG-EE provides the functionality to define CPG-specific decision logic rules based on the decision variables in the CPG and to execute the rules based on patient clinical data to provide CPG based recommendations. The CPG-EE comprises two main modules: (i) Rule Authoring Module and (ii) Rule Execution Module.

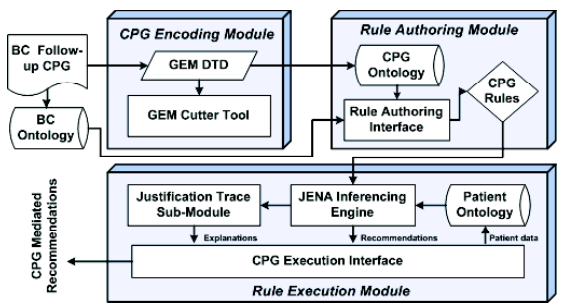


Figure 2 - System diagram of the CPG-EE

Rule Authoring Module

The Rule Authoring Module provides users an interface to specify decision logic rules, using a native CPG rule syntax, based on the decision logic inherent within the CPG. The rule authoring process is guided by the knowledge, relationships and constraints represented within the CPG ontology and the domain ontology (in our case the BC ontology). In this way, rule authoring is constrained by pre-defined knowledge and hence ensures the sanity of the decision rules.

The CPG ontology is designed to semantically model the structure of a CPG in order to annotate the decision variables and logic structures inherent within the CPG. Our CPG ontology is based on the GEM DTD which provides a characterization of different elements of a CPG. In particular we utilized the knowledge component of the GEM DTD and mapped it to a *Recommendation* class that entails the procedural, conditional or imperative knowledge of the CPG. The decision and action variables that constitute the premises and conclusions of a CPG rule, respectively, are explicitly stated in the CPG ontology, and these variables are utilized in authoring CPG rules. In the CPG ontology, the decision variables are represented as a sub-class. For execution purposes we added a new property variable.name to the decision.variable, such that its value is derived from all properties in the Domain Ontology.

Rule Authoring is performed by defining decision rules in the logic tag of CPG ontology as follows: *Step 1:* Select decision variables from the Decision Variable List, which represents the body (premises) of the rule; *Step 2:* Select the action variable from the Decision Variable List, which represents the head (conclusion) of the rule; *Step 3:* For each decision and action variable in the rule, an equality/ inequality relation is defined with either a variable, a value, a binary algebraic formula, another decision variable or list of decision variables. We give a rule authoring example, where we assign the variable names i.e. properties to *decision variables* (coded as dv and each with a unique #) as well as action variables (coded as *av* and each with a unique #). In case of a rule

> **IF dv1** i.e., *Patient_is_on_medication* = Tamoxifen (property of class Patient_Type) **AND dv2** i.e. *Rx_apply_to_recommendation* = ? (property of class Rx_Recommended)
>
> **THEN a1** i.e., *Patient_is_recommended.* (property of class Patient_Type) = **dv5**

The derivation for this rule is as follows. The Patient_Type_1 which is an instance of the class Patient_Type is *on medication*, Tamoxifen. Patient_Type_1 is the resource for this rule. The treatment i.e. Tamoxifen is an instance of class Treatment, which has a property *apply to recommendation*, whose value is 'query about vaginal bleeding'. Since we have specified in the rule that the value for *a1* (*Patient_is_recommended*) is same as the value for *dv2* (*Rx_apply_to_recommendation*), which according to the BC ontology is 'query about vaginal bleeding', the recommendation for this patient type is to query about vaginal bleeding.

Rule Execution Module

The Rule Execution Module executes the CPG rules based on a patient instance to infer CPG based recommendations. Rule execution is performed by a logic-based inference engine—i.e., JENA. The processing of this module is as follows: (i) The CPG rules are transformed from their native syntax to JENA rule syntax; (ii) The patient data is acquired through the CPG execution interface (see Figure 4) to form an instance of a patient, based on the Patient Ontology, that incorporates patient properties such as *age, gender, medical history*, etc. The values of the patient properties serve as input to the execution engine; (iii) The JENA inference engine uses the CPG rules and the patient instance to build an inference model using backward chain reasoning. The outcome is a set of inferred recommendations based on the patient data; (iv) A justification trace of the inferred recommendations is generated to explain the reasons for the proposed recommendations.

BCF-DSS in Action

We present an example to demonstrate BCF-DSS in action. The clinical case is: *"A BC patient who is overweight, complains of fatigue and is experiencing vaginal bleeding in the absence of obvious cause. She has a family history of osteoporosis and is on Aromatase inhibitors. She also wishes to get pregnant and wants to know whether this is a viable option".*

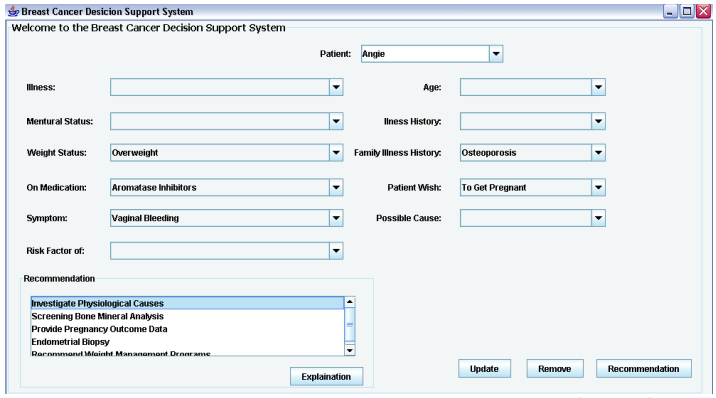


Figure 3 - The CPG execution interface for BCS-DSS, used to collect the patient data and to give recommendations

The family physician records the patient's properties using the BCC-DSS user interface (shown in Figure 3). The physician presses the *Recommendation* button and is provided five recommendations (shown in the bottom left box). The physician can seek an explanation for any recommendation by highlighting it and pressing the *Explanation* button. Figure 4 shows the explanation interface that includes the CPG description (upper left box) for the recommendation, the reasons for the proposed recommendation (upper right box) and the related references (lower middle box); all explanation material is derived from the annotated BC CPG. Finally, the justification trace (see Figure 5) for the inferred recommendations is as follows:

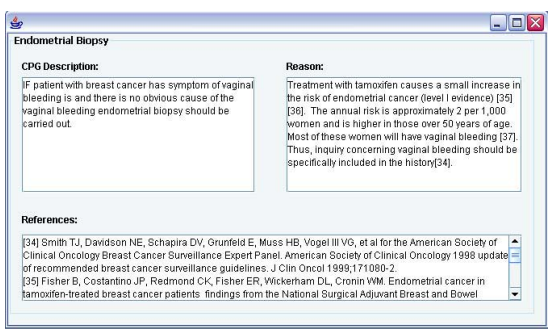


Figure 4 - The explanation interface of the BCF-DSS

Angie → has weight status = Overweight
Overweight → associated with recomm = Recommend Weight Management Programs
Angie → is Recommended = *Recommend Weight Management Programs*

Angie → has symptom = Vaginal Bleeding
Vaginal Bleeding → is not cause by = Menstruation or Obvious Cause
Menstruation or Obvious Cause → ms apply to diagnostic test = Endometrial Biopsy
Angie → is Recommended = *Endometrial Biopsy*

Angie → has family history of Illness = Osteoporosis
Osteoporosis → illness apply to recomm = Treatment Should Include Bisphosphonates
Angie → is Recommended = *Treatment Should Include Bisphosphonates*

Angie → has wish = To Get Pregnant
To Get Pregnant → associated with recommendation = Provide Pregnancy Outcome Data
Angie → is Recommended = *Provide Pregnancy Outcome Data*

Angie → is on medication = Aromatase Inhibitors
Aromatase Inhibitors → Rx apply to Diagnostic Test = Screening Bone Mineral Analysis
Angie → is Recommended = *Screening Bone Mineral Analysis*

Angie → has symptom = Fatigue
Fatigue → apply to recommendation = Investigate Physiological Causes
Angie → is Recommended = *Investigate Physiological Causes*

Figure 5 - Justification trace for the recommendations

Concluding remarks

We have developed a CPG based interactive clinical decision support system for the BC follow-up to be used in the primary care setting. Our approach is innovative since we have linked the CPG ontology to the breast cancer domain ontology from which rules were derived. This approach can also be applied to CPG in other medical specialties. The objective of this project is to promote knowledge translation to primary care settings in Nova Scotia so that family physicians can take on the responsibility for the BC follow-up care, thereby reducing the strain on specialist cancer centers within Nova Scotia. The project also aims to create an interactive environment for family physicians to facilitate customized patient management and educational information for an individual patient.

Acknowledgement

This project is funded by Nova Scotia Health Research Foundation (Canada), through an innovation research grant. Special thanks to Dr. Grunfeld for the BC-CPG and her support.

References

[1] Pryor DB, Fortin DF. Managing the delivery of health care: care-plans/managed care/practice guidelines. *Int J Biomed Comput.* 1995;39:105–9.

[2] Zielstorff RD. Online Practice Guidelines, Issues, Obstacles and Future Practice, *JAMIA*. 1998;5:227–236.

[3] Lomas J, Anderson GM, Domnick-Pierre K. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. *N Engl J Med.* 1989; 321:1306–1311.

[4] Greco PJ, Eisenberg JM. Changing physicians' practices. *N Engl J Med.* 1993;329:1271–1273.

[5] Sanders GD, Nease RF, Jr., Owens DK. Design and Pilot Evaluation of a System to Develop Computer-based Site-specific Practice Guidelines from Decision Models; *Med Decis Making* 2000;20:145–159.

[6] Grunfeld E, Levine M, Julian J, Prtichard K, Coyle D, Mirsky D, Verma S, Dent S, Sawka C, Ginsburg D. A randomized controlled trial of routine follow-up for early stage breast cancer: a comparison of primary care versus specialist care. *J Clin Oncol* 2004; 22:43

[7] Grunfeld E, Dhesy-Thind S, Levine M. Guidelines for the Care and Treatment of Breast Cancer: Follow-up after treatment for breast cancer (2005 update), *Can. Med. Assoc. J.*, May 2005; 172: 1319 - 1320

Address for correspondence

Syed Sibte Raza Abidi Email: sraza@cs.dal.ca

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Learning Causal and Predictive Clinical Practice Guidelines from Data

Subramani Mani[a], Constantin Aliferis[a], Shanthi Krishnaswami[b,c] and Theodore Kotchen[c]

subramani.mani@vanderbilt.edu constantin.aliferis@vanderbilt.edu shanthi.krishnas@vanderbilt.edu tkotchen@mcw.edu

[a] *Department of Biomedical Informatics, Vanderbilt University, Nashville TN 37232*
[b] *Department of Rheumatology, Vanderbilt University, Nashville TN*
[c] *Department of Medicine, Medical College of Wisconsin, Milwaukee WI*

Abstract

Clinical practice guidelines (CPG) propose preventive, diagnostic and treatment strategies based on the best available evidence. CPG enable practice of evidence-based medicine and bring about standardization of healthcare delivery in a given hospital, region, country or the whole world. This study explores generation of guidelines from data using machine learning, causal discovery methods and the domain of high blood pressure as an example.

Keywords:

clinical practice guidelines, machine learning, prediction, compliance, causal discovery, high blood pressure

Introduction and background

Clinical practice guidelines (CPG) have been developed to streamline practice of medicine in a complex healthcare environment. Guidelines are the foundation of evidence-based medicine (EBM) and provider compliance with evidence-based CPG facilitates the practice of evidence-based medicine.

Typically guidelines are developed by a committee of domain experts specifically constituted for the purpose. After a rigorous review and analysis of published evidence over a period of time, the committee publishes its set of guidelines. However, this process is laborious, time consuming and expensive. See for the example the joint national committee report on high blood pressure [1].

Based on published evaluations of CPG, Grimshaw and Russell showed that practice guidelines streamlined the process of care and contributed to better outcomes in general [2]. There has been considerable interest in guideline representation for making them computable. Many different frameworks for guideline implementation such as PRODIGY [3], Arden Syntax [4], GLIF [5], PROforma [6], Asbru [7] and EON [8] have been proposed.

The fields of machine learning (ML), data mining and knowledge discovery including methods for learning cause and effect relationships have matured over the years with applications to clinical and biological data. There have only been limited attempts at using machine learning and knowledge discovery methods for generating practice guidelines from data [9-12]. The goal of Abston et al. [9] was to discover the variation between the American College of Cardiology guidelines for management of acute myocardial infarction and documented practices in a tertiary care facility in Salt Lake city. The study did not address generation of new guidelines from data. Mani et al. [10] presented a two-stage machine learning model as a data mining method to develop clinical practice guidelines and showed its value in dementia staging. It modeled the methodology used by clinicians by deriving intermediate concepts in the first phase and using the intermediate concepts for dementia staging. However, it is not clear if the method is generalizable across different domains. Morik et al. [11] used a combination of prior knowledge from experts and learning from data for clinical protocol generation and validation. Sboner et al. [12] used machine learning techniques to model the decision making of dermatologists in melanoma diagnosis. Svatek et al. [13] describe a data mining approach based on association rule learning for checking guideline adherence. However, the clinical validity of the approach could not be ascertained due to the small sample size of the study.

Machine learning techniques have typically been explored for disease screening [14], differential diagnosis [15], and other outcome measures [16]. A method based on inductive logic programming for learning qualitative physiological models from clinical data has also been described [17]. However, to the best of our knowledge causal discovery methods have not been explored for guideline generation from data.

This paper explores automated generation and compliance checking of guidelines. Guidelines can be broadly classified as predictive guidelines or prevention/intervention guidelines based on their goals. Note that predictive guidelines are sufficient for diagnosis (diagnose disease D1 from symptoms S1 and S2). For prevention and intervention we need a cause and effect interpretation for the guidelines.

We now introduce a formal notion of causality. We define the *causal influence* of a variable A on a variable B using the *manipulation criterion* [18, 19]. The manipulation criterion states that if we had a way of setting just the values of A and then measuring B, the causal influence of A on B will be reflected as a change in the conditional distribution of B. That is, there exist values a_1 and a_2 of A such that $P(B \mid \text{set } A = a_1)$ $P(\boldsymbol{B} \mid \text{set } A = a_2)$. A causal influence of

variable *A* on variable *B* is represented as an arc from *A* to *B* i.e. *A* *B*. We say that variable *A* causally influences variable *B* if and only if *A* and *B* satisfy the manipulation criterion.

A causal influence of a variable *A* on a variable *B* is said to be *unconfounded* if and only if there is *no* measured or unmeasured variable *C* that is a common cause of variables *A* and *B*.

Materials and methods

Algorithms

For learning predictive guidelines we selected two machine learning algorithms with the following properties.

1. Perform classification (prediction) tasks well.
2. The generated models are comprehensible to humans.
3. The models can be easily implemented as computerized guidelines.

The C4.5 algorithm that uses the decision tree representation formalism [20] and RIPPER [21] that has the format of an *If ... Then rule* were selected. Decision trees and rules generate clear descriptions of how the ML method arrives at a particular classification.

For checking the cause and effect interpretation of the guideline we used the FCI algorithm [18]. The FCI algorithm takes as input a dataset D and outputs a graphical model consisting of edges between them that have a cause and effect interpretation. The FCI algorithm can handle hidden (unmeasured) variables and sample selection bias that are likely to be present in real-world datasets. There are other causal discovery algorithms (for example, PC [18]) that output a causal Bayesian network (CBN) model [22, 23], incorporating all the variables represented in a dataset. However, PC assumes that all the variables in a domain are observed and there are no unobserved variables. There are also causal discovery algorithms that take a local approach and output causal relationships of the form "variable *A* causally influences variable *B*". LCD [24] and BLCD [25] are two such algorithms. The local causal discovery algorithms are particularly suitable for large datasets.

In this study we apply C4.5, RIPPER and FCI to the high blood pressure (HBP) dataset that is described below.

Dataset

The prevlanence of high blood pressure in the US is approaching 30% and the rate of prevalence is also showing an increasing trend [26]. The dataset used in this work is part of an ongoing NIH funded study with Dr. Kotchen, T as the principal investigator and its goal is to ascertain genetic determinants and other causal factors of high blood pressure. The HBP dataset is a population based dataset consisting of data collected from consenting African Americans between the ages of 18 and 55 years in Milwaukee and neighboring areas. Anthropometric measurements included height, weight, waist, hip, arm circumferences and skinfold thickness measured at different sites. Consenting subjects who satisfied inclusion and exclusion criteria were admitted for a 2 day inpatient protocol to obtain additional hemodynamic and renal measurements under standardized controlled conditions. Exclusion criteria included secondary hypertension, diabetes, creatinine >2.2 mg/dl, body mass index (BMI) >35, recent stroke or myocardial infarction, malignancy and substance abuse including alcohol. Currently 369 people are enrolled (202 hypertensives and 167 normotensives). 47.3% of the normotensives and 53% of hypertensives were females. The average blood pressure of normotensive subjects was 114/74 vs. 147/96 in hypertensive subjects. 31% of the subjects were on antihypertensive medication. We selected 23 variables after excluding patient identifiers and redundant variables (variables derived from other variables present in the dataset). See Table 1 for the list of variables selected. The variables were categorized based on either the established risk levels of each variable for cardiovascular diseases or the study-specific cutpoints using the 90th / 10th percentile levels of values in normotensive subjects. The outcome (class) variable was coded H for hypertensive and N for normotensive based on the following guideline. If the outpatient blood pressure (OP-BP) was greater than or equal to 140/90 or the subject was on BP medication (BP-MED), the outcome variable (HBP) was coded H, otherwise it was coded N. All the independent variables were categorized as 0 and 1 (one representing risk for high blood pressure). Thus the guideline used for creating the outcome variable was our target hypothesis for learning which is given below.

If OP-BP = 1 or BP-MED = 1, HBP = H; **else** HBP = N.

We created two datasets DS1 and DS2 as follows. DS1 had 21 variables after excluding OP-BP and BP-MED that were used for generating the outcome HBP. The purpose of creating DS1 was to ascertain the causal factors for the outcome variable as a baseline before generating guidelines. DS2 included all the 23 variables. A third dataset DS3 was created from DS2 by toggling the value of the outcome variable for a randomly selected 10% of the subjects i.e. if the value was H it was changed to N and vice versa. This was to artificially create a set of 37 patients for whom the guideline was violated. DS3 was used for verification of guideline compliance. A fourth dataset DS4 was also created with data on this set of 37 patients for whom the class label was manipulated. C4.5 and Ripper were run using datasets DS2, DS3 and DS4. FCI was run on datasets DS1 and DS2.

The FCI program is available from the Tetrad project site at Carnegie Mellon (www.phil.cmu.edu/tetrad/tetrad4.html). We used the Java implementation of C4.5 and RIPPER available in the Weka machine learning software package [27]. We performed a ten fold cross-validation for C4.5 and RIPPER and report the results based on the test cases that were not used in model building. Note that for causal discovery cross-validation is not relevant because we are not performing classification or regression.

Table 1: Variables used in the study

Plasma Aldo/ plasma Renin ratio
Age
Cardiac output baseline
Creatinine clearance
Gender
High density lipoprotein
Heart rate at baseline
Insulin Resistance
Potassium excretion
Calculated Low density lipoprotein
Sodium excretion
Baseline renal Blood flow
Baseline Renal vascular resistance
Systemic Vascular resistance Index baseline
Stroke Volume baseline
Serum Triglycerides
Urine 24hrs Microalbumin
Waist circumference risk
Outpatient Hypertension yes/no
Glucose risk
Screening Height
Outpatient high BP /Normal BP
On antihypertensive medication

Results

Figures 1 and 2 present the results of the application of C4.5 and Ripper to the HBP DS2 dataset.

```
OP-BP = 0
|   BP-MED = 0: N
|   BP-MED = 1: H
OP-BP = 1: H

Number of Leaves  : 3
```

Figure 1 - A C4.5 tree from DS2

The C4.5 tree classified all the 369 instances correctly with an accuracy of 100%. The precision and recall were 1 for the H and N classes.

RIPPER

DS2 Ripper rules:

```
(OP-BP = 0) and (BP-MED = 0) => HBP=N
 => HBP = H
```

Figure 2 - A RIPPER rule set from DS2

The RIPPER classified all the 369 instances correctly with an accuracy of 100%.

Results of C4.5 and Ripper on DS3

Both C4.5 and Ripper misclassified the 37 instances for which the class assignments had been changed. All the other instances were classified correctly. Both the algorithms generated the same models shown in Figures 1 and 2.

Results of C4.5 and Ripper on DS4

The results of application of C4.5 and Ripper to the DS4 dataset are shown in Figure 3 and Figure 4 respectively. Recall that the DS4 dataset has just the 37 instances with labels manipulated.

```
OP-BP = 0
|   BP-MED = 0: H
|   BP-MED = 1: N
OP-BP = 1: N

Number of Leaves  : 3
```

Figure 3 - A C4.5 tree from DS4

The DS4 C4.5 tree classified 35 out of the 37 instances correctly with an accuracy of 95%. The precision for class N was 1 and recall for class H was 1.

RIPPER

DS4 Ripper rules:

```
(OP-BP = 0) and (BP-MED = 0) => HBP=H
 => HBP = N
```

Figure 4 - A RIPPER rule set from DS4

The RIPPER rule set classified 35 out of the 37 instances correctly with an accuracy of 95%.

FCI results

When applied to the HBP DS1 dataset, FCI output four possible causal factors for high blood pressure. Table 2 enumerates them. Note that when the relationship is categorized as "O>", there could be a common cause or a feedback loop.

Discussion

The machine learning algorithms C4.5 and Ripper recovered the study guideline that was used for assigning labels to the outcome variable when the guideline was followed in all the cases (DS2). The same guideline was also generated from the dataset when the guideline was not followed in 10% of the subjects (DS3). The subjects for whom the guideline was not followed were also identified as misclassified instances by C4.5 and Ripper. This shows that ML methods can be used for generating simple guidelines and guideline compliance checking.

Table 2 - The output of FCI from DS1

Aldo/Renin ratio O> HBP
Age O> HBP
Renal vascular resistance <> HBP
Waist risk <> HBP

A O> B Means there is definitely an arrowhead at B.
There may or may not be an arrowhead at A.
A <> B denotes that there is a common cause for A and B.

Table 3 - The output of FCI from DS2

OP-BP HBP
BP-MED HBP
HDL Cholesterol Insulin resistance
Insulin resistance Glucose risk

The results shown in Figure 3 and Figure 4 using the dataset DS4 with the 37 misclassified cases from DS3 shows the alternate guideline model generated for those cases. This is the alternate guideline that was followed for these instances. This shows that when a specified guideline is not followed, ML can be used to identify any alternate guideline that might have been used instead. Note though that the instances would be misclassified only if the application of the alternate guideline changed the outcome variable.

We also used a causal discovery algorithm (FCI) to ascertain the cause and effect basis of the guideline. Using DS1 FCI output four causal influences shown in Table 2. When DS2 was used OP-BP and BP-MED were output by FCI as unconfounded causal factors for HBP. Note that OP-BP and BP-MED are causal factors for HBP based on the manipulation criterion for causality. However, they are also causal based on the definition of the BP study guideline. Two other unconfounded causal relationships from the domain were also output by FCI (see Table 3).

Decision tree models and If … Then rules are expressive and easily interpretable by humans. Moreover, the tree and rule formats are also suitable for computerized guidelines and hence useful for incorporation as decision support tools in electronic medical record systems.

Our study addresses the question of generating new practice guidelines in a data driven way and explores the role of causal discovery along with traditional machine learning approaches for guideline generation from data.

Two relevant issues that come up in guideline application are generalizability and customization. A guideline developed in one institution or organization may not be exactly applicable in another practice setting. Likewise, a guideline developed by a committee of national or international experts might need to be customized to a local setting. Fridsma et al. have developed a knowledge-based approach to customization based on the separation of site specific and site independent factors that can be identified from the knowledge of the organization and understanding of its workflow [28]. We believe that data driven machine learning approaches could be a useful tool in the overall effort to make guidelines generalizable and customizable.

Limitations

The dataset that we used in this study was from a population-based study of high blood pressure. There were only a small number of variables in the dataset. Only two machine learning and one causal discovery algorithm were used in this study. The guidelines were also very simple.

Conclusions and future work

In this paper we presented a machine learning approach to generate guidelines from data, check for guideline compliance and if non-compliant for a set of patients, generate the alternate guideline used. We also provided a method for ascertaining whether the guideline has a causal semantics using a causal discovery algorithm.

In future we plan to apply machine learning and causal discovery algorithms to different medical datasets involving more complex guidelines for further evaluation of our approach.

References

[1] Chobanian, A, Bakris, G, Black, H, Cushman, W, Green, L, Izzo, J, Jones, D, Materson, B, Oparil, S, Wright, J, Roccella, E and NHBPEPCC. Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. Hypertension l2003;42: 1206-1252.

[2] Grimshaw, JM and Russell, IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet l1993;342: 1317-22.

[3] Purves, I, Sugden, B, Booth, N and Sowerby, M. The PRODIGY project--the iterative development of the release one model. In: AMIA fall symposium; 1999. p. 359-63.

[4] Clayton, P, Pryor, T, Wigertz, O and Hripcsak, G. Issues and structures for sharing knowledge among decision-making systems: The 1989 Arden Homestead Retreat. In: K. LC, editor. Thirteenth Annual Symposium on Computer Applications in Medical Care: IEEE Computer Society Press; 1989. p. 116-21.

[5] Ohno-Machado, L, Gennari, J, Murphy, S, Jain, N, Tu, S, Oliver, D, Pattison-Gordon, E, Greenes, R, Shortliffe, E and Barnett, G. The guideline interchange format: a model for representing guidelines. JAMIA l1998;5: 357-72.

[6] Fox, J, Johns, N and Rahmanzadeh, A. Disseminating medical knowledge: The PROforma approach. Artiff Intell Med l1998;14: 157-81.

[7] Miksch, S, Shahar, Y and Johnson, P. Asbru: a task-specific, intention-based, and time-oriented language for representing skeletal plans. In: Seventh Workshop on Knowledge Engineering Methods and Languages (KEML-97): Milton Keynes, UK; 1997.

[8] Musen, M, Tu, S, Das, A and Shahar, Y. EON: a component-based approach to automation protocol-directed therapy. JAMIA l1996;3: 367-88.
[9] Abston, K, Pryor, T, Haug, P and Anderson, J. Inducing practice guidelines from a hospital database. In: JAMIA Supplement; 1997. p. 168--172.
[10] Mani, S, Shankle, WR, Dick, MB and Pazzani, MJ. Two-Stage Machine Learning Model for Guideline Development. In: Artificial Intelligence in Medicine; 1998.
[11] Morik, K, Imboff, M, Brockhausen, P, Joachims, T and Gather, U. Knowledge discovery and knowledge validation in intensive care. Artiff Intell Med l2000;19: 225-249.
[12] Sboner, A and Aliferis, CF. Modeling clinical judgment and implicit guideline compliance in the diagnosis of melanomas using machine learning. In: AMIA fall symposium; 2005. p. 664-668.
[13] Svatek, V, Riha, A, Peleska, J and Rauch, J. Analysis of guideline compliance---a data mining approach. In: K. Kaiser, S. Miksch and S. W. Tu, editors. Symposium on Computerized Guidelines and Protocols: IOS Press; 2004. p. 157-161.
[14] Shankle, W, Mani, S, Pazzani, M and Smyth, P. Detecting Very Early Stages of Dementia from Normal Aging with Machine Learning Methods. In: E. Keravnou, C. Garbay, R. Baud and J. Wyatt, editors. Lecture Notes in Artificial Intelligence: Artificial Intelligence in Medicine, AIME97: Springer; 1997. p. 73--85.
[15] Ohmann, C, Yang, Q, Moustakis, V, Lang, K and Elk, van PJ. Machine Learning Techniques Applied to the Diagnosis of Acute Abdominal Pain. In: P. Barahona and M. Stefanelli, editors. Lecture Notes in Artificial Intelligence: Artificial Intelligence in Medicine, AIME95: Springer; 1995. p. 276--281.
[16] Cooper, G, Aliferis, C and Ambrosino, R. An evaluation of machine learning methods for predicting pneumonia mortality. Artificial Intelligence in Medicine l1997;9: 107--138.
[17] Hau, DT and Coiera, EW. Learning qualitative models of dynamic systems. Machine Learning l1993;26: 177-211.
[18] Spirtes, P, Glymour, C and Scheines, R. Causation, Prediction, and Search. In. Cambridge, MA: MIT Press; 2000.
[19] Cooper, GF. An overview of the representation and discovery of causal relationships using Bayesian networks. In: C. Glymour and G. F. Cooper, editors. Computation, Causation, and Discovery. Cambridge, MA: MIT Press; 1999. p. 3--62.
[20] Quinlan, J. C4.5: Programs for Machine Learning. In. Los Altos, California: Morgan Kaufmann; 1993.
[21] Cohen, WW. Fast effective rule induction. In: ICML; 1995. p. 115-123.
[22] Pearl, J. Probabilistic Reasoning in Intelligent Systems. In. San Francisco, California: Morgan Kaufmann; 1991.
[23] Neapolitan, RE. Probabilistic Reasoning in Expert Systems: Theory and Algorithms. In. New York: John Wiley and Sons; 1990.
[24] Cooper, GF. A Simple Constraint-Based Algorithm for Efficiently Mining Observational Databases for Causal Relationships. Data Mining and Knowledge Discovery l1997;1: 203--224.
[25] Mani, S. A Bayesian Local Causal Discovery Framework (doctoral dissertation). In: Intelligent Systems Program: University of Pittsburgh; 2005.
[26] Hajjar, I and Kotchen, TA. Trends in prevalence, awareness, treatment, and control of hypertension in the United States, 1988 - 2000. JAMA l2003;290: 199-206.
[27] Witten, IH and Frank, E. Data Mining: Practical machine learning tools and techniques. In. 2nd ed: Morgan Kaufmann, San Francisco; 2005.
[28] Fridsma, DB, Gennari, J and Musen, M. Making generic guidelines site-specific. In: AMIA fall symposium; 1996.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Supporting Therapy Selection in Computerized Clinical Guidelines by Means of Decision Theory

Stefania Montani, Paolo Terenziani, Alessio Bottrighi

DI, Universiy of Piemonte Orientale,Alessandria, Italy

Abstract

Supporting therapy selection is a fundamental task for a system for the computerized management of clinical guidelines (GL). The goal is particularly critical when no alternative is really better than the others, from a strictly clinical viewpoint. In these cases, decision theory appears to be a very suitable means to provide advice. In this paper, we describe how algorithms for calculating utility, and for evaluating the optimal policy, can be exploited to fit the GL management context.

Keywords:

Decision Theory, Clinical Guidelines

Introduction

Clinical guidelines (GL) can be defined as a means for specifying the "best" clinical procedures and for standardizing them. In recent years, the medical community has started to recognize that a computer-based treatment of GL provides relevant advantages, such as automatic connection to the patient databases and, more interestingly, decision making facilities; thus, many different approaches and projects have been developed to this hand (see e.g. [1,2]). As a matter of fact, decision making is a central issue in clinical practice. In particular, supporting therapy selection is a critical objective to be achieved. Consider that, when implementing a GL, a physician can be faced with a choice among different therapeutic alternatives, and identifying the most suitable one is often not straightforward. Unlike clinical protocols [3], that specify what is the only admissible procedure in a given situation, GL are used in domains in which different choices are actually possible (see the example in the results section). Alternatives can be pruned relying both on *site-related* contextual information (e.g. due to the unavailability of certain resources in a given hospital), and on *patient-related* contextual information (due to the peculiarities of the single patient on which the GL is being applied). The works in [4, 5] show how decision models can be resorted to in order to help physicians in defining and contextualizing GL. However, also when the GL has been properly contextualized, it is frequent to find more then one option left, and sometimes no one of these remaining alternatives is really "better" than the others, from a strictly clinical viewpoint. In clinical practice, various selection parameters (such as the costs and the effectiveness of the different procedures) can be available to physicians when executing a GL. The computer-based GL systems described in the literature offer sophisticated formalizations of these decision criteria. Nevertheless, the available information is often only qualitative in nature, and "local" to the decision at hand: it does not take into account the consequences of the choice, in terms of actions to be implemented, and future decisions to be taken along the path stemming from the selected alternative. On the other hand, the possibility of obtaining a complete scenario of the decision consequences, considering the probability of the different therapy outcomes, the utilities associated to the different health states, and the money, time and resources spent, would be clearly an added value for physicians and hospital administrators. Decision theory seems a natural candidate as a methodology for covering this task. To this hand, a systematic analysis of the main GL representation primitives, and of how they could be related to decision theory concepts has been recently proposed [6]. Since, at a sufficiently abstract level, the GL representation primitives treated in that work are shared by all the systems in the literature [7], that contribution can be seen as the first step towards the implementation of a tool within any of the such approaches. In this paper, we start from such knowledge representation results, to describe how decision theory algorithms (to calculate utility and to obtain the optimal policy) can be exploited, when the goal is the one of supporting therapy selection in a GL management system. We also analyse complex situation which may arise due to the presence of certain types of control flow relations among GL actions, namely iterations and parallel executions. A practical application of this work is represented by the tool which is being implemented in GLARE, a domain-independent system for GL acquisition and execution [8, 9]. The algorithmic choices, and the technical issues discussed in this paper will therefore refer to this specific example. In particular, an earlier version of GLARE already embedded a facility able to calculate costs, time and resources required to complete paths in a GL (details can be found in [9], and are briefly sketched in the next section); the decision theory support can be seen as an extension of that work. The paper is structured as follows: in the next section we clarify the goal of our decision theory tool and we summarize the previous results in the direction of supporting therapy selection, i.e. the concept

mapping work and the main features of the GLARE's cost-collection facility. In the results section we describe technical issues about the implementation (referring to the specific example of GLARE). Finally the last section addresses some concluding remarks.

Materials and methods

Designing the main features

We envision the possibility of adopting a decision theory tool for supporting therapy selection in two fashions. First, it can operate in the on-line modality, when applying the GL actions one at a time to the patient at hand, by automatically retrieving the patient's data from the Hospital Information System (HIS). In this case, at the time at which a therapeutic decision has to be taken, the facility is able to provide local pros and cons of the various alternatives. Secondly - and more interestingly - the tool can be used off-line, if the physician wants to make a simulation of the consequences of a therapeutic alternative, by evaluating the patient's evolution along the different paths stemming from the decision at hand, typically until the end of the GL is reached. The possibility of collecting this global information is crucial to allow her making a well informed choice. This modality would be useful also for education purposes. From the algorithmic viewpoint, this working mode generalizes the first one; therefore, in the rest of the paper, we will concentrate on the off-line modality. Off-line simulation typically involves a series of temporally consequent decisions. The clinical GL can therefore be seen as a dynamic decision problem. In particular, as we will briefly motivate below, the GL can be mapped onto a (completely observable) Markov Decision Process (MDP), in which the sequence of therapeutic decisions generates the sequence of (patient) states. The typical goal of a decision theory tool is to find the optimal policy, i.e. the sequence of decisions able to maximize the expected utility. In the context of GL, it is possible to adapt and simplify this task by limiting the decisions to be considered to therapy selections among clinically equivalent alternatives (see the example in the results section); we will call them non-trivial decisions henceforth. Moreover, we propose to realize an implementation in which also costs, resources and time spent to complete any path in the GL can be obtained, and can be coupled with the calculation of the expected utility along the path itself. Note that fixing the path means fixing the policy that has to be applied, i.e. knowing which alternative will be chosen at any decision. Finally, the user should be always allowed to select the part of the GL s/he wants to focus on (as a default, the overall GL will be taken into account). The possibility of selecting only a portion of the GL seems to us particularly relevant, since we aim at supporting only non-trivial decisions, while the GL will typically include paths where therapeutic choices do not require the adoption of the decision theory facility; moreover, concentrating only on a subpart of the GL will obviously reduce the computation time. All the technical details about a concrete implementation of a decision theory tool within the system GLARE are described in the results section.

Previous work

Concept mapping

In [6], a knowledge representation contribution, aimed at mapping the GL primitives to decision theory concepts, was provided. In particular, at a sufficiently abstract level, GL representation formalisms share the following assumptions (for the terminology used here, we refer in particular to [8, 10]). First, a GL can be represented as a graph, where nodes are the *actions* to be executed, and arcs are the *control relations* linking them. It is possible to distinguish between *atomic* and *composite* actions (plans), which can be defined in terms of their atomic components via the *has-part* relation. Three different types of *atomic actions* can then be identified: (1) *work actions*, i.e. actions that describe a procedure which must be executed at a given point of the guideline; (2) *query actions*, i.e. requests of information from the outside world; (3) *decision actions*, used to model the selection among different alternatives. Decision actions can be further subdivided in *diagnostic decisions*, used to make explicit the identification of the disease the patient is suffering from, and *therapeutic decisions*, used to represent the choice of a path in the GL, containing the implementation of a particular therapeutic process (henceforth, we will concentrate on (non-trivial) therapeutic decisions, that we want to support). *Control* relations establish which actions can be executed next, and in what order. For example, actions could be executed in *sequence*, or in *parallel*. Moreover, the *alternative* relation describes how alternative paths can stem from a decision action, and the *repetition* relation states that an action has to be repeated several times (maybe a number of times not known a priori, until a certain *exit condition* becomes true). In a well-formed GL, a decision action is preceded by a query action, that is adopted to collect all the patient's parameters necessary (and sufficient) for taking the decision itself. Each decision is therefore based on an (explicit or implicit) data collection completed at decision time, and does not depend on the previous history of the patient. We can thus say that the GL describes a discrete-time first-order Markov model, since each time a query action is implemented, the patient's situation is completely re-assessed, and an (explicit or implicit) query action is always found before a decision action. This observation justifies the mapping of GL primitives to the field of decision theory, and in particular allows us to represent a GL as a MDP. Despite the fact that Markov processes have been often applied to medical domains (e.g. for disease management in [11]), the difficulties in relying on them to simulate clinical processes are well known. In particular, the resulting model can be very demanding not only with regard to the amount of data needed to specify the probability distribution underlying the stochastic process, but also computationally [12]. Nevertheless, these limitations appear to be less critical in the domain of clinical guidelines, where rather strict design policies are typically applied. Therefore, when dealing with GL, some simplifications hold. In particular, as already observed, a first-

order Markov model is sufficient to capture the GL dynamics. Moreover, the process modelled by the GL is completely observable, since in a GL a decision can be taken only if all the required parameters have been collected: if some needed data are missing, the query action will wait for them and the decision will be delayed. It is then straightforward to define the *state* as the set of patient's parameters that are normally measured for taking decisions and for assessing therapy outcomes. Query actions are the means for observing the state. *State transitions* are produced by all the work actions between two consecutive non-trivial therapeutic decisions. Finally, the *utility* of a state can be evaluated in terms of life expectancy, corrected by Quality Adjusted Life Years (QALYs) [13].

The cost-collection facility

GLARE already incorporates a decision support facility, able to assist physicians in choosing among therapeutic alternatives [9]. Relying on this tool, it is possible to compare different paths in the GL, by simulating what could happen if a certain choice was made. In particular, users are helped in calculating the "cost" of the paths themselves, in order to select the cheapest choice. Costs are not interpreted just as monetary expenses, but also as resources and time required to complete GL actions. Note that, when running the tool, if a composite action is found, it is expanded in its components, and the reasoning facility is recursively applied to each of them, by analysing all the decision actions that appear at the various decomposition levels. At the end of this process, the tool displays the values of the collected parameters (costs, resources, times) gathered along each path. The final decision is then left to the physician.

Results

Within GLARE, the facility described in the previous section is being extended, by allowing: (1) the identification of the optimal policy, and (2) the calculation of the expected utility along a path. In order to implement these functionalities, we had to take into account the following issues.

Focusing

As already observed, the possibility of selecting only a sub-part of a given GL is a fundamental issue to be addressed, since it allows one to skip the paths on which decision theory support is not required. In our tool, path selection has been conceived as the first step of the interaction with the user. Technically speaking, the mechanism works as follows: through a user-friendly graphical interface, the physician is asked to indicate the starting node (normally the decision at hand) of the paths to be compared and (optionally) the ending nodes (otherwise all possible paths exiting the starting node will be taken into consideration, until the end of the GL). For every decision action within each path, s/he is allowed to restrict to a subset of alternatives. Moreover, the selection process is recursively applied to composite actions. All the paths pruned by this procedure will be ignored by the subsequent steps of the reasoning process (i.e. mapping to the Markov model and extraction of the optimal policy, or calculation of the expected utilities).

Parallel actions

In case two or more composite actions, each one containing non-trivial decisions, have to be executed in parallel along a selected path, the mapping towards the corresponding Markov model (needed to provide functionality 1 above) is not straightforward. As a matter of fact, in this situation the order of execution of the various actions is not univocally provided. The policy we have chosen to adopt to this hand is the one of calculating just one possible order of execution, compatible with the available temporal constraints [14], and to rely on it.

Generating the Markov model

Once path selection and parallel actions management have been addressed, the next step towards the calculation of the optimal policy (see functionality 1 above) is the mapping of the GL to the corresponding Markov model. The algorithm, which relies on the concept mapping results described in the previous section, automatically produces the conversion by operating as follows: for every selected path, in correspondence to a non-trivial decision (or to an exit point of the GL) it generates a new state (unless the same state was already identified). It then collects all the work actions between two consecutive states, and builds a macro-action; the macro-action determines the transition between the two states with a given probability. Probability values are extracted from the medical literature - when possible; otherwise, in the current implementation, they can be obtained from interviews with expert physicians. Note that, for those medical fields in which the medical literature does not provide these numbers, it is reasonable to expect this information to be available in the near future. As a matter of fact, the increasing exploitation of HIS and of computerized GL management tools will allow for the collection of large amounts of clinical practice data, on which it will be easy to draw statistics, at least at the local level. Consider also that relying on local data is not necessarily a limitation: remember that a guideline always needs to be contextualized to the features of the hospital in which it has to be implemented, before its exploitation begins [15].

Repeated actions

On the Markov model, classical algorithms for evaluating the optimal policy can be relied on [16, 17]. When dealing with a finite time horizon, the dynamic programming algorithm can be easily applied [16]. Nevertheless, in the context of clinical GL, it is not infrequent that the number of states, though finite, is not known a priori. This situation may be induced by the presence of iterations to be repeated several times, until a certain exit condition becomes true. The number of repetitions could therefore vary among different executions of the same GL. To handle this problem, the choice made in GLARE is the one of relying on algorithms for calculating the optimal policy on an infinite time horizon (as a matter of fact, "infinite" can be used in

the meaning of "unknown a priori"), and in particular on value iteration [17].

Simplifications

Algorithms can be simplified in case the required output is not the optimal policy, but the expected utility of all the different paths selected by the user (see functionality 2 above). Note that following a path corresponds to apply a specific policy, i.e. to make the hypothesis of knowing what is the decision to be taken at any decision node. In case of a finite time horizon, since the policy is known, we can calculate the utility by applying the dynamic programming algorithm, avoiding to maximize the expected value of the cumulative utility function with respect to the different actions. The corresponding costs can be summed up resorting to the functionality described in the previous section. On the other hand, when iterations with an exit condition have to be tackled, we can ask the user physician the minimum and maximum number of times that the action has to be repeated, given the specific patient's characteristics. Then, we can generate the paths corresponding to these two extreme situations, thus reducing to a finite time horizon, on which it is possible to calculate utility and costs as described above. In the case of utility, however, we can also keep working on an infinite time horizon, not having to rely on the physician's estimate of the number of iterations, and resort to value iteration. Again, since the policy is fixed, the algorithm can be simplified by avoiding maximization wrt the different actions. This second strategy is clearly not applicable to costs, which are additive. It is up to the user physician to select the preferred output in these cases. As an example, we present an application of the GLARE decision theory tool to a GL for asthma treatment. Figure 1 shows part of the GL. Patients affected by mild persistent asthma (see upper branch in the figure) may be treated by four different therapies (as indicated by the four edges exiting the T1 node): inhaled beta-2-agonists (A), oral beta-2-agonists (OA), inhaled anticholinergics (IA) or theophylline (TH). Each therapy implementation consists of a daily dose. Basically, the four drugs for mild asthma are clinically equivalent; therefore, indications about implications of each alternative could be useful in deciding. If the therapy does not work, the patient could worsen to moderate asthma. In this case, another therapeutic decision has to be taken (action T2), in order to implement a more effective treatment. It is possible to select between inhaled steroids (S) and inhaled steroids plus bronchodilators (SB). Again, these drugs have to be provided daily. Periodically (e.g. weekly), the patient's state is re-assessed, and the therapeutic decision has to be repeated in a loop, until this guideline becomes not applicable for the patient at hand (because asthma is now severe, and a different guideline has to be referred to, or because asthma improves: this case is not explicitly represented in figure 1). In the GLARE formalism, the treatment of mild asthma has to be represented as an iterated plan with an exit condition (see the previous section), that corresponds to the onset of moderate asthma. Analogous considerations hold for the treatment of moderate asthma. The utility of the mild persistent asthma state is 86, while the utility of moderate asthma is 82 and the one of severe asthma is 78 [18]. Table 1 lists the probabilities of transition, that were provided by medical experts (see acknowledgements). Since the GL includes two cycles, which are repeated a number of times not known a priori, in this example the optimal policy has been calculated using the value iteration algorithm, which has identified A as the optimal therapy in case of mild asthma, and SB as the optimal therapy in case of moderate asthma. We have applied (simplified) value iteration also for calculating the utility of all the possible paths in the GL; the corresponding values are reported in table 2.

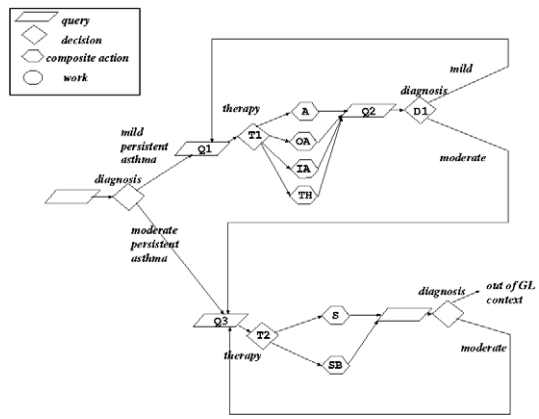


Figure 1 - Part of the asthma treatment guideline. Node types legend is provided as well.

Table 1 - Probabilities of transition between states.

Therapy	From	To	Prob.
A	Mild	Mild	0.9
A	Mild	Moderate	0.1
OA	Mild	Mild	0.85
OA	Mild	Moderate	0.15
IA	Mild	Mild	0.6
IA	Mild	Moderate	0.4
TH	Mild	Mild	0.5
TH	Mild	Moderate	0.5
S	Moderate	Moderate	0.85
S	Moderate	Severe (out of GL)	0.15
SB	Moderate	Moderate	0.95

Therapy	From	To	Prob.
SB	Moderate	Severe (out of GL)	0.05

Conclusions

In the context of clinical GL, it is not infrequent to identify actions of therapeutic selection which would benefit from a decision theory support. Embedding a decision theory facility would therefore be an added value for a computerized system for GL management. In this paper, we have discussed the main features that should characterize such a tool. In particular, we have identified what simplifications can be made with respect to classical decision theory approaches, meant to be applied to a generic domain. Moreover, we have underlined what integrations and specific choices are required, on the other hand, to deal with non trivial GL control flow constructs, namely iterations and parallel executions. These considerations appear to be general enough to be exploited when designing a decision theory tool within any of the systems described in the literature. As an example, we have presented here the features of the system GLARE's facility. We believe that a tool developed along these lines would be able to provide a valuable support to physicians, thus reinforcing the claim that the adoption of AI techniques can provide relevant advantages in the (semi)-automatic treatment of clinical GL, and favouring the actual exploitation of computer science facilities within the medical community.

Table 2 - Utilities of the asthma GL states having fixed the various policies (i.e. paths). Numbers have been normalized with respect to the maximum value.

Policy	Utility Mild asthma	Utility Moderate
A-S	0.57590	0.24231
A-SB	1	0.66641
OA-S	0.46470	0.24231
OA-SB	0.88879	0.66640
IA-S	0.32570	0.24230
IA-SB	0.74979	0.66639
TH-S	0.30902	0.24230
TH-SB	0.73311	0.66639
A-S	0.57590	0.24231

Acknowledgments

Information about utilities and state transition probabilities in the example were obtained from Prof. G. Molino and Prof. M. Torchio of Azienda Ospedaliera S. Giovanni Battista in Turin, with whose cooperation the implementation work is being carried on. This research has been partially supported by a grant from Regione Piemonte.

References

[1] Fridsma D.B. (Guest Ed.) Special issue on workflow management and clinical guidelines. Journal of the American Medical Informatics Association, 1(22), 1-80,2001.

[2] Gordon C. and Christensen J.P. editors. Health Telematics for Clinical Guidelines and Protocols. IOS Press, Amsterdam, 1995.

[3] Purves I., Computerised guidelines in primary health care: Reflections and implications, Telematics for Clinical Guidelines and Protocols, C.Gordon and J.P. Christensen, 57-74, IOS Press, Amsterdam, 1995

[4] Sanders G.D., Nease R.F., Owens D.K., Design and pilot evaluation of a system to develop computer-based site-specific practice guidelines from decision models, Medical Decision Making 20, 145-149, 2000

[5] Sanders G.D., Nease R.F., Owens D.K., Publishing web-based guidelines using interactive decision models, Journal of Evaluation in Clinical Practice, 20, 145-159, 2001

[6] Montani S., Terenziani P., Exploiting decision theory concepts within clinical guideline systems: towards a general approach, International Journal of Intelligent Systems 21(2006) 585-599

[7] Peleg M. and al., Comparing models of decision and action for guideline-based decision support: a case-study approach, JAMIA, ed. American Medical Informatics Association, 10, 52-68, 2003

[8] Terenziani P., Molino G., Torchio M., A modular approach for representing and executing clinical guidelines, Artificial Intelligence in Medicine, 23(3), 249-276, 2001

[9] Terenziani P., Montani S., Bottrighi A., Torchio M., Molino G., Supporting physicians in taking decisions in clinical guidelines: the GLARE "what-if" facility, Journal of the American Medical Informatics Association (JAMIA) Symposium supplement, 772-776, 2002

[10] Fox J., Johns N., Rahmanzadeh A., Thomson R., Disseminating medical knowledge: the PROforma approach, Artificial Intelligence in Medicine, 14(1-2), 157-182, 1998

[11] Lucas P, de Bruijn N., Schurink K, Hoepelman A., A probabilistic and decision-theoretic approach to the management of infectious disease at the ICU, Artificial Intelligence in Medicine 19 (2000) 251–279

[12] Boutilier C., Dean T., Hans S., Decision-theoretic planning: structural assumptions and computational leverage. J Artif Intell Rev 1999;11:1–94.

[13] Gold M.R., Siegel J.E., Russell L.B., Weinstein M. C., Cost-Effectiveness in Health and Medicine, Oxford University Press, New York, 1996

[14] Vila L., A survey on temporal reasoning in artificial intelligence, AI Communications, 1(7), 4-28, 1994

[15] Terenziani P., Montani S., Bottrighi A., Torchio M., Molino G., Correndo G., A context-adaptable approach to clinical guidelines, Proc. Medinfo (World Congress for Health Informatics) 2004, San Francisco, September 2004, M. Fieschi et al. eds., IOS Press, Amsterdam, 169-173

[16] Bellman R., Dynamic programming, Princeton University Press, 1957

[17] Tijms H.C., Stochastic modelling and analysis: a computational approach, Wiley and Sons, 1986

[18] Moy M.L., Israel E., Weiss S.T., Juniper E.F., Dube L., Drazen J.M., Clinical predictors of health-related quality of life depend on asthma severity, Am. J. Respir. Crit. Care Med., 163(4), 924-929, 2001

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Development, Deployment and Usability of a Point-of-Care Decision Support System for Chronic Disease Management Using the Recently-Approved HL7 Decision Support Service Standard

David F. Lobach, MD, PhD, MS[a]; Kensaku Kawamoto, PhD[a]; Kevin J. Anstrom, PhD[b]; Michael L. Russell, MD[c]; Peter Woods[c]; and Dwight Smith[c]

[a]Division of Clinical Informatics, Department of Community and Family Medicine; [b]Department of Biostatistics and Bioinformatics; and [c]Duke Health Technology Solutions; Duke University, Durham, North Carolina, USA

Abstract

Clinical decision support is recognized as one potential remedy for the growing crisis in healthcare quality in the United States and other industrialized nations. While decision support systems have been shown to improve care quality and reduce errors, these systems are not widely available. This lack of availability arises in part because most decision support systems are not portable or scalable. The Health Level 7 international standard development organization recently adopted a draft standard known as the Decision Support Service standard to facilitate the implementation of clinical decision support systems using software services. In this paper, we report the first implementation of a clinical decision support system using this new standard. This system provides point-of-care chronic disease management for diabetes and other conditions and is deployed throughout a large regional health system. We also report process measures and usability data concerning the system. Use of the Decision Support Service standard provides a portable and scalable approach to clinical decision support that could facilitate the more extensive use of decision support systems.

Keywords:

clinical decision support, web services, service-oriented architecture, decision support service, Health Level 7

Introduction

The volume of clinical knowledge reported in the biomedical literature is rapidly increasing [1]. However, the dissemination of this medical knowledge through traditional channels (e.g., publication of evidence-based guidelines, continuing medical education) has been found to be inadequate [2]. In the United States, a recent nationwide audit assessing 439 quality indicators found that American adults receive only about half of recommended care [3], and the U.S. Institute of Medicine has estimated that up to 98,000 Americans die each year as the result of preventable medical errors [4]. Similar deficiencies in care quality have been found in other industrialized nations [5].

Selected for best paper award.

In seeking to address this crisis in care quality, one of the most promising strategies for optimizing patient care and ensuring patient safety involves the use of computer systems that support clinical decision making [6]. Such clinical decision support (CDS) systems represent one of the most effective means for improving clinician compliance with evidence-based care standards [7]. The utilization of CDS systems, however, remains limited in most healthcare facilities [8]. While multiple factors have contributed to this limited adoption of CDS systems, one important factor has been the lack of an efficient method for encapsulating, processing, and delivering executable medical knowledge for use in clinical software applications [9].

In attempting to overcome the difficulty of re-using medical knowledge encoded in a machine-executable format, knowledge engineers have generally taken two approaches [10]. As one approach, systems such as PRODIGY [11], SAGE [12], and First DataBank's Drug Information Framework™ [13] provide access to their executable knowledge base using standard application programming interfaces. As a second approach, methods including GLIF3 [14], GEM [15], and Arden Syntax [16] encode knowledge using a common formalism, so that encoded rules can be consistently interpreted by system-specific interpreters. Despite these significant efforts, a dominant framework has not emerged for sharing executable medical knowledge, due in part to the following challenges. First, some formalisms, such as the Drug Information Framework™ [13], focus on specific knowledge domains and are not extendable to other domains. Second, many formalisms are designed for use in specific types of CDS applications and are difficult to adapt for use in other types of applications. Third, many formalisms are difficult to understand due to their conceptual complexity. Finally, many existing methods require significant investments in infrastructure, such as a system-specific compiler [17].

With regard to actual systems, most CDS systems are designed for specific institutional settings. Thus, despite repeated validation of their effectiveness, the utilization of CDS systems remains the exception rather than the rule in most practices across the country. This underutilization is

largely because CDS systems with demonstrated efficacy have generally been designed as an extension to a specific electronic health record (EHR) system and cannot be easily transferred to clinical sites that use a different EHR or no EHR at all [18]. Thus, while several CDS systems have validated efficacy [7], they have not provided a general, widely implementable solution to the problem of substandard care.

Given this portability problem, several CDS systems have been developed as stand-alone systems that operate independent of existing information systems [19, 20]. However, the lack of system-to-system integration introduces new problems. Stand-alone systems require substantial and continuous data entry to maintain the current and complete patient information required for the generation of accurate care recommendations. In addition, many stand-alone systems require clinicians to proactively access the system. The reliance on such proactive usage is particularly troublesome when the system only needs to be accessed for a subset of patients. As a result, stand-alone systems oftentimes fail to become a part of routine workflow, which is critical to a CDS system's ability to improve clinical practice [7, 18].

We have developed a CDS Web service that overcomes many of these challenges known as the SEBASTIAN Decision Support Service (DSS) [21]. A few other sites have also used a service-based approach to decision support [22, 23]. The service interface of the SEBASTIAN DSS has provided the foundation of a new HL7 draft standard known as the HL7 Decision Support Service standard, which was formally adopted as a draft HL7 standard during the September 2006 ballot cycle [24].

In this paper, we describe the first production use of a decision support system based on the new HL7 DSS standard. We also report on process measures related to the use of the system, the usability of the application enabled by the DSS, and lessons learned from this experience. The results of our efforts can assist others in using the HL7 DSS in other settings.

Materials and methods

SEBASTIAN DSS

We have previously developed and described the SEBASTIAN (an acronym for System for Evidence-Based Advice through Simultaneous Transaction with an Intelligent Agent across a Network) DSS, which enables machine-executable medical knowledge to be re-used across applications and institutions [21]. The SEBASTIAN DSS is implemented as a Web service, in which software functionality is provided over the Internet and extensible markup language (XML) messages are used to communicate with client systems [25].

Point-of-care chronic disease management system enabled by SEBASTIAN DSS

Implementation setting. The SEBASTIAN DSS was used to implement a point-of-care chronic disease management system within the Duke University Health System, located in Durham, North Carolina. The health system handles over 60,000 hospitalizations and 1.2 million outpatient encounters a year.

System functionality. The chronic disease management system was added as a tab in the patient summary section of the Duke electronic record viewing system, known as the eBrowser. The system currently provides care recommendations for diabetes management. The system is currently being enhanced to support preventive health reminders and disease management for hypertension, asthma, and dyslipidemia.

System architecture. The system architecture of the chronic disease management system is shown in Figure 1. For simplicity's sake, this figure and the accompanying text description focus solely on how the system generates care recommendations for diabetes management. Care recommendations for health maintenance and for other chronic medical conditions will be generated in a manner that parallels the information flow described below.

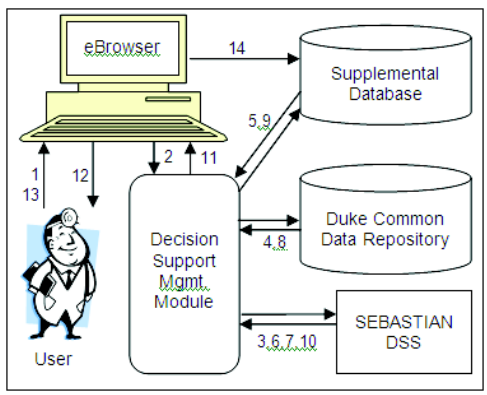


Figure 1 - Architecture of Duke Chronic Disease Management System

The information flow in the chronic disease management system is coordinated by a Decision Support Management Module. This module interacts with the SEBASTIAN DSS in order to identify the data required for evaluating a patient and to obtain machine-interpretable conclusions regarding a patient. This module also interacts with two patient data sources to retrieve the data required by SEBASTIAN.

According to this design, the clinician enters an identification number for a specific patient into the Duke eBrowser (arrow #1). The eBrowser passes this number to the Decision Support Management Module (#2). This module then contacts the SEBASTIAN DSS to find out what data are needed to determine if this particular patient has diabetes mellitus (#3). The management module then obtains the required data from the primary clinical database, known as the Duke Common Data Repository (CDR) (#4), as well as from a supplemental database that contains data pertinent to chronic disease management but not captured in the CDR (#5). The management module then sends these data to the SEBASTIAN DSS to determine if the

patient does have diabetes and receives back a result indicating whether or not the patient is diabetic (#6). If the patient is diabetic, the management module then inquires of the SEBASTIAN DSS what data are needed to run the diabetes management rules and receives back the list of required data (#7). Next, the management module obtains the required data from the CDR and the supplemental database (#8, 9). The module then sends these data to the SEBASTIAN DSS and receives back the recommendations for the chronic management of diabetes (#10). The management module formats these data and posts the results to the eBrowser (#11), where they are then viewed by the clinician (#12). If the clinician has additional information related to the diabetes recommendations, such as knowledge that an influenza immunization was given at a local pharmacy, she can then enter this data into the data entry screen on the eBrowser (#13). These data are then stored in the supplemental database (#14).

System development. The chronic disease management system was implemented through collaboration between the operational Duke Health Technology Solutions (DHTS) group and the academic Division of Clinical Informatics. The Decision Support Management Module was written in C#, and this module was interfaced with the eBrowser to fulfill clinicians' requests for disease management recommendations. The availability of the chronic disease management functionality was announced in a "What's New" pop-up window along with other eBrowser enhancements when the DSS first became available.

Evaluation

Process measures. The number of distinct system users and the number of distinct patients for whom the system was used were collected weekly. System performance was evaluated in terms of the time required for generating diabetes care recommendations for a set of representative patients.

Usability survey. We surveyed users of the SEBASTIAN system at a family medicine primary care practice within the Duke University Health System. The survey was conducted after the trial system had been in use for three months, but prior to release of the final system. The surveyed population consisted of attending family medicine physicians, family medicine residents, physician extenders, a pharmacist, and a dietitian.

The usability survey was adapted from validated survey instruments for measuring end-users' computing satisfaction [26] and their perceptions of system usefulness and ease of use [27]. Questions from these surveys were grouped into constructs related to content (precise information needed, sufficient information); accuracy; format (clear, useful); usefulness (improves job performance, increases productivity, makes job easier); and ease of use (easy to use, flexible, understandable). These questions utilized a Likert scale in which 3 is a neutral value on a 1-to-5 scale. Confidence intervals and p-values were calculated using large-sample Z tests and a null hypothesis stating that the population mean is 3. The survey respondents were also asked to provide free-text comments regarding the system.

Results

System implementation

The SEBASTIAN DSS has been implemented as a production system to support chronic disease management for the entire Duke University Health System. On August 10, 2006, the trial version of the system was made available to the 1509 attending physicians and 832 residents within the health system, as well as to physician extenders and to ancillary personnel with access to the clinical components of the eBrowser. The production version of the system was released on November 13, 2006 (Figure 2). The current system provides support for diabetes management. As discussed earlier, the system is also being expanded to provide support for health maintenance and for multiple other chronic medical conditions.

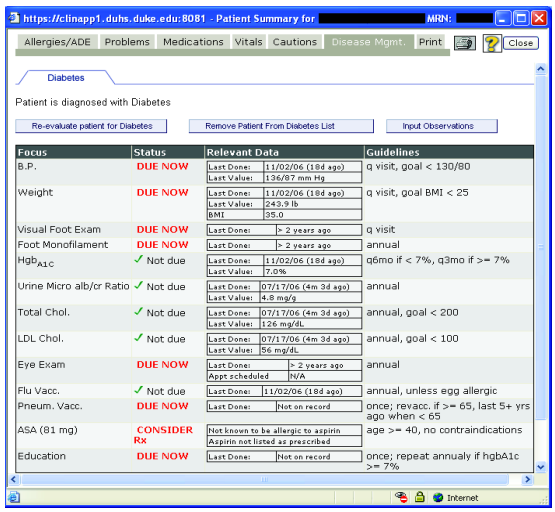


Figure 2 - Screenshot from Duke Chronic Disease Management System showing recommendations for a patient with diabetes

Process measures

Since inception, with relatively minimal promotion, the system has been accessed by 758 unique users. This user community is comprised of physicians, physician assistants, nurse practitioners, medical and allied health students, and ancillary service providers. Weekly, since inception, the system is used on average by 85 distinct users (range 72 to 104) to access chronic disease management recommendations for diabetes on an average on 176 distinct patients (range 113 to 225).

With regard to system performance, the system requires an average of 10.9 seconds to return care recommendations after a request is submitted through the eBrowser. Most of this time is spent retrieving the necessary data from the common data repository. Interactions with the SEBASTIAN DSS account for less than a second of the overall processing time.

Usability survey

Twenty (63%) of the 32 available providers responded to the usability survey. Of these 20 providers, 17 were active users of the chronic disease management system. The results

from the usability surveys are depicted in Figure 3. User responses for content, accuracy, format, and ease of use were significantly favorable ($p < 0.005$). Responses for the usefulness construct was not statistically different from the neutral response ($p = 0.76$). An analysis of users' free-text comments regarding the system revealed that this neutral perception of system usefulness arose in large part from the perception that the system took too much time to use.

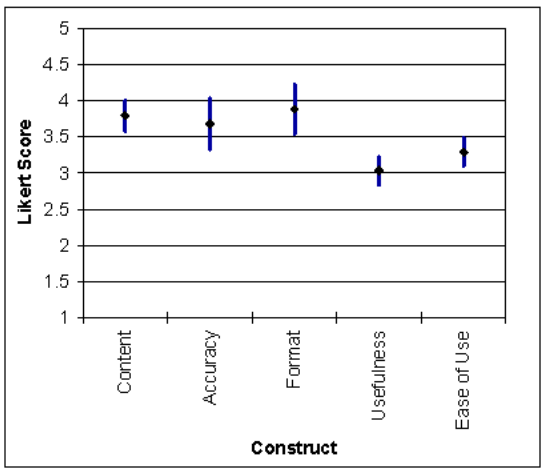


Figure 3 - Results of usability survey, presented as mean scores on usability constructs with associated 95% confidence intervals

Discussion

In this paper, we have described the first implementation of a decision support system based on the recently approved HL7 Decision Support Service draft standard. This implementation validates the operational usefulness of this new HL7 draft standard in the context of a large health system.

The SEBASTIAN DSS approach to meeting a client's CDS needs offers many important strengths and overcomes many of the limitations of other approaches to implementing CDS systems. As one strength, this approach provides a clear mechanism for various clinical information systems to leverage the SEBASTIAN CDS knowledge base. Second, the functionality and interface of the SEBASTIAN DSS is the basis of the HL7 DSS standard. As a result, SEBASTIAN is aligned with an emerging industry consensus on how to deliver CDS using a service-based approach. Third, our approach could support CDS needs at a regional or national level. Fourth, our approach is aligned with the *Roadmap for National Action on CDS* commissioned by the Office of the National Coordinator for Health IT, which calls for a service-based approach to the management and delivery of CDS content [28]. Fifth, our approach is scalable and standards-based. Finally, our approach has been vetted and validated through its production-level use in multiple settings, including in the Duke University Health System for point-of-care chronic disease management and in the North Carolina Medicaid program for population health management [21].

A limitation to our approach is that some clients may be hesitant to have a third party host a critical clinical application. However, the increasing success of vendors that provide electronic health record systems using an application service provider model points to the fact that this hesitation can be overcome with a track record of reliable service. Second, our approach has not yet been validated for several important types of CDS applications (e.g. computerized physician order entry systems). Finally, we do not yet have documented evidence that using a DSS will lead to outcomes desired by a client, such as improved performance on care quality and pay-for-performance metrics. However, there is strong evidence that CDS systems implemented in the manner described in this manuscript reliably produce significant improvements in clinical practice [7]. Also, we are currently conducting several evaluation studies to assess the impact of SEBASTIAN-enabled CDS systems on patient care.

From the limited survey results based on use of the prototype system, we conclude that the clinicians perceived the SEBASTIAN disease management system as easy-to-use, accurate, and appropriate in content and format. The main objection to the system was in the area of usefulness, because of the additional time required for obtaining the care recommendations from the system. We believe this objection was valid, as the prototype system sometimes took over twenty seconds to return the diabetes care recommendations. As noted earlier, this delay resulted from the time required to retrieve the required data from the Duke clinical data repository. In recognition of this limitation, indexing changes were made to the clinical data repository, so that the system currently takes approximately ten seconds to generate a care recommendation summary. We are in the process of introducing additional performance-enhancing strategies, including multi-threaded data retrieval and pre-caching of patient data, in order to reduce the time required for obtaining the care recommendations to under five seconds in the short-term and under one second in the long-term.

In terms of lessons learned, we have observed that the time required to extract the necessary data in order to run the decision support rules may be excessive for busy clinicians. As a consequence, we recommend the retrieval and caching of necessary data prior to the actual request to activate the DSS. This implementation has also taught us that the usefulness of a DSS could be enhanced by the availability of software libraries that facilitate a client's interaction with a DSS. Our implementation experience has also shown us the need to accommodate local configuration of the DSS. Many clinical areas lack a clear consensus for how care should be delivered, and a DSS needs to be able to allow for such "gray" areas of knowledge. We also identified the need to allow a DSS to be configured for individual patients, by allowing rules to be inactivated when clinically inappropriate or declined in deference to patient preference. Finally, we have come to recognize common patterns of CDS needs faced by client applications (e.g., need to determine if patient has condition X and is need of test Y). Recognition of these patterns may allow for more generalized approaches to DSS-client interactions.

Conclusion

This paper is the first report of a decision support application that is built upon the HL7 Decision Support Service draft standard. This application provides disease management information to clinicians for diabetes and other conditions at the point of care. The implementation of this disease management application in a large health system validates the concept of a Decision Support Service.

Acknowledgements

The authors wish to thank Ms. Sherri Gerringer for assistance with survey data collection and management. This work was funded in part by grant R01 HS015057 from the Agency for Healthcare Research and Quality and by grant R41 LM009051 from the National Library of Medicine.

Disclosure: Drs. Lobach and Kawamoto are part owners of Kedasys, LLC, a holding company that has a pending patent application for the intellectual property related to the SEBASTIAN approach to instantiating a Decision Support Service. Through prior agreement, SEBASTIAN is available for use within the Duke University Health System at no cost.

References

[1] National Library of Medicine. MEDLINE Fact Sheet. Available at: http://www.nlm.nih.gov/pubs/factsheets/medline.html. Accessed 12/4/06.

[2] Davis DA, Thomson MA, Oxman AD, and Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. AMA 1995:274:700-705.

[3] McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, and Kerr EA. The quality of health care delivered to adults in the United States. N Engl J Med 2003:348:2635-2645.

[4] Kohn LT, Corrigan JM, and Donaldson MS, eds. To Err is Human: Building a Safer Health System. Washington, DC: National Academy Press, 1999.

[5] Vincent C, Neale G, and Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. BMJ 2001:322:517-519.

[6] Hunt DL, Haynes RB, Hanna SE, and Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. JAMA 1998:280:1339-1346.

[7] Kawamoto K, Houlihan CA, Balas EA, and Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005. 330:765-768.

[8] Southern California Evidence-based Practice Center. Costs and Benefits of Health Information Technology. Evidence Report/Technology Assessment, Number 132. Agency for Healthcare Research and Quality. April 2006. http://www.ahrq.gov/downloads/pub/evidence/pdf/hitsyscosts/hitsys.pdf. Accessed 11/28/06.

[9] Kawamoto K. Integration of knowledge resources into applications to enable clinical decision support: architectural considerations. In Greenes RA, ed. Medical Decision-Making: Computer-Based Approaches to Achieving Healthcare Quality and Safety. Burlington, MA: Elsevier, Inc. 2007; pp. 503-538.

[10] Boxwala AA, Tu S, Peleg M, Zeng Q, Ogunyemi O, Greenes RA, Shortliffe EH, and Patel VL. Toward a representation format for sharable clinical guidelines. J Biomed Inform 2001:34:157-169.

[11] Johnson PD, Tu S, Booth N, Sugden B, and Purves IN. Using scenarios in chronic disease management guidelines for primary care. Proc AMIA Symp 2000:389-393.

[12] Tu SW, Musen MA, Shankar R, Campbell J, Hrabak K, McClay J, Huff SM, McClure R, Parker C, Rocha R, Abarbanel R, Beard N, Glasgow J, Mansfield G, Ram P, Ye Q, Mays E, Weida T, Chute CG, McDonald K, Molu D, Nyman MA, Scheitel S, Solbrig H, Zill DA, and Goldstein MK. Modeling guidelines for integration into clinical workflow. Medinfo 2004:2004:174-178.

[13] First DataBank. Drug Information Framework Information Page. Available at: http://www.firstdatabank.com/products/di_framework/. Accessed 12/4/06.

[14] Peleg M, Boxwala AA, Ogunyemi O, Zeng Q, Tu S, Lacson R, Bernstam E, Ash N, Mork P, Ohno-Machado L, Shortliffe EH, and Greenes RA. GLIF3: the evolution of a guideline representation format. Proc AMIA Symp 2000:645-649.

[15] Shiffman RN, Michel G, Essaihi A, and Thornquist E. Bridging the guideline implementation gap: a systematic, document-centered approach to guideline implementation. J Am Med Inform Assoc 2004:11:418-426.

[16] Pryor TA, and Hripcsak G. The Arden syntax for medical logic modules. Int J Clin Monit Comput 1993:10:215-224.

[17] Karadimas HC, Chailloleau C, Hemery F, Simonnet J, and Lepage E. Arden/J: an architecture for MLM execution on the Java platform. J Am Med Inform Assoc 2002:9:359-368.

[18] Zielstorff RD. Online practice guidelines: issues, obstacles, and future prospects. J Am Med Inform Assoc 1998:5:227-236.

[19] Hunt DL, Haynes RB, Hayward RS, Pim MA, and Horsman J. Patient-specific evidence-based care recommendations for diabetes mellitus: development and initial clinic experience with a computerized decision support system. Int J Med Inf 1998:51:127-135.

[20] Henry SB, Douglas K, Galzagorry G, Lahey A, and Holzemer WL. A template-based approach to support utilization of clinical practice guidelines within an electronic health record. J Am Med Inform Assoc 1998:5:237-244.

[21] Kawamoto K, and Lobach DF. Design, implementation, use, and preliminary evaluation of SEBASTIAN, a standards-based Web service for clinical decision support. Proc AMIA Symp 2005:380-384.

[22] Huang C, Noirot LA, Heard KM, Reichley RM, Dunagan C, and Bailey TC. Implementation of Virtual Medical Record Object Model for a Standards-Based Clinical Decision Support Rule Engine. Proc AMIA Symp 2006:958.

[23] Goldberg HS, Vashevko M, Pastilnik A, Smith K, Plaks N, and Blumenfeld BM. Evaluation of a Commercial Rule Engine as a Basis for a Clinical Decision Support Service. Proc AMIA Symp 2006:294-298.

[24] Health Level 7. HL7 Decision Support Service specification (Draft Standard for Trial Use). Available at: http://www.hl7.org/v3ballot/html/infrastructure/dss/dss.htm. Accessed 11/27/06.

[25] Cerami E. Top ten FAQs for Web services. Available at: http://webservices.xml.com/pub/a/ws/2002/02/12/webservicefaqs.html. Accessed 11/27/06.

[26] Doll WJ, and Torkzadeh G. The Measurement of End-User Computing Satisfaction. MIS Quarterly 1988:12:259-274.

[27] Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989:13:319-340.

[28] Kawamoto K, and Lobach DF. Proposal for fulfilling strategic objectives of the U.S. roadmap for national action on decision support through a service-oriented architecture leveraging HL7 services. J Am Med Inform Assoc 2007:14:146-155.

Address for correspondence

David F. Lobach. email:- david.lobach@duke.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

The TAR Model: Use of Therapeutic State Transitions for Quality Assurance Reporting in Chronic Disease Management

Gaikwad R[1], Warren J[1], Kenealy T[2]

[1]*Department of Computer Science, The University of Auckland, New Zealand*
[2]*Department of General Practice and Primary Health Care, The University of Auckland, New Zealand*

Abstract

Chronic disease management represents one of the challenges for health informatics and demands the appropriate application of information technology for improved patient care. This paper presents an approach to quality assurance reporting wherein the recommendations of evidence-based clinical practice guidelines are considered in the context of empirical therapeutic state-transitions (in terms of changes in individual patient prescriptions over time). We apply a Transition-based Audit Report (TAR) model to antihypertensive prescribing and related data as stored in a New Zealand General Practice Management System database. The results provide a set of quality indicators and specific patient cohorts for potential practice quality improvement with strong linkage to the selected guidelines and observed practice patterns. We see the TAR model primarily as a tool to enable internal quality improvement efforts, but also to be of relevance for focusing pay-for-performance programs.

Keywords:

hypertension, medical audit, practice guidelines, quality assurance, quality indicators

Introduction

General medical practices in New Zealand have a 99% rate of using purpose-built Practice Management Systems (PMS) software, with 89.7% using the PMS to record prescriptions and 71.8% reporting that their General Practitioners (GPs) use the PMS to record full clinical notes [1]. This high rate of PMS uptake is likely to have a positive influence on quality of care, since electronic prescribing systems have been shown to reduce the frequency of medication errors [2]; although, conversely, such systems have also been associated with specific types of increased risk in patient care [3]. The ubiquity of PMS use by GPs presents a distinct opportunity to use data mining to extract empirical information about chronic disease management practices, particularly the well-recorded prescribing practices, such as prescription of antihypertensive drugs. Cardiovascular Disease (CVD) is the number one killer in New Zealand [4], so the motivation to improve the management of related risk factors is, of course, considerable. Understanding actual practice in antihypertensive prescribing serves to inform our understanding of CVD risk management.

Therapeutic state transitions are points in time when the status of key aspects of a patient's therapy change [5]. Analysis of therapeutic state transitions in General Practice – notably, the pattern of GPs' prescribing acts for individual patients over time – shows promise as the basis for development of high-specificity interactive decision support alerts [6]. In this paper we examine the potential for therapeutic state transitions to inform the development of an audit report, making it a Transition-based Audit Report (TAR). We examine the TAR model in the context of antihypertensive prescribing using the PMS data from a medium-size metropolitan General Practice in New Zealand.

Methodology

This work utilizes the approach of therapeutic state transition based analysis as set forth by Warren et al. [5]. We examine how therapy (in this case, antihypertensive prescribing, qualified by coded observations and diagnoses) changes over time for the individual patients that make up the cohort of interest. In this paper we focus on a TAR model of report formulation to support internal practice quality improvement efforts, and to inform broader discussion of appropriate performance metrics.

Setting and data

Data was extracted from the PMS of a metropolitan New Zealand General Practice which employs about four full-time equivalent GPs and two full-time equivalent nurses, belongs to the large Primary Health Organization in the country, and is sited in an area that is socio-economically and ethnically diverse. The investigation related to all prescribing undertaken by the practice between January 1, 2005 and September 25, 2006. Analysed PMS data included age, gender, prescriptions, "classifications" (problem codes, largely as Read Codes), and observations including blood pressure (BP), serum creatinine and serum potassium measurements. Characteristics of the data extract are given in Table 1. Identity of patients was blinded to the analysts (RG and JW), but could be recombined for follow-up by the GP author (TK, one of the main GPs of the practice) for purposes of follow up in patient

care. As such, the orientation of the investigation was on internal practice quality assurance and improvement with a secondary goal of methodology development. Some data cleaning was necessary; notably, to compensate for mis-punctuation and trailing zeros in entry of BP measurements (which can be accomplished through any of multiple pathways in the user interface of the PMS, including via free-text notes). About 98% of BP measurements were usable.

Table 1 - Characteristics of data extract

General description of data extract (investigated period: Jan 1, 2005 - Sept 25, 2006)	No. of records
Patients[#]	14835
Encounter	84393
Classifications	
Total	8704
Distinct patients classified	3383
Hypertension alone or as comorbidity	481 (23%)
Hypertension without diabetes mellitus	396 (82%)
Hypertension with diabetes mellitus	85 (18%)
Prescriptions	
Total	59473
Antihypertensive prescriptions*	8466 (14%)
Measurements - Blood pressure	10418

[#] distinct patients in data extract classified irrespective of being hypertensive or not * prescriptions with antihypertensive agents

Choice of guidelines and therapeutic state variables

Within the scope of this paper, the focus was on the quality of antihypertensive therapy (AHT) prescribing, thus emphasizing CVD risk management and appropriate variation for comorbidities, notably diabetes. We chose clinical practice guidelines on the criteria that they were evidence-based, applicable to the clinical practice in the area of study and current [7, 8, 9, 10]. Generic and/ or brand names for AHT agents were identified from the data and classified by referring to [11]. There were 44 generic and/ or brand names of drugs identified as belonging to six AHT drug groups (see Table 2). The therapeutic drug groups (admittedly broad groupings) were given compact labels – A, B1, B2, B3, C and D – based on their general order in [7]. These six groups identify our therapeutic state variables for analysis; that is, changes in the presence or absence of prescription of drugs from these groups defines therapeutic state transitions.

Therapeutic state transitions

The process of prescribing by the GP produces two events; one marking the start of the prescription (therapy) and the second, an implicit event marking the expected end of the prescription if directions given by the physician have been properly adhered to. The instructions (*signatura*) given by the GP consist of dosing, frequency, repeats (refills) and a duration (computed by the PMS) per prescription. We identified 290 distinct *signatura* used by the GPs for the drugs of interest in the data extract. The PMS-computed duration was accurate except where the GP had overridden default dosing instructions, wherein we computed a corrected duration; e.g. "take 10mg od" of Plendil ER was interpreted as a directedconsumption of 4*(2.5mg Tab Plendil ER)*1 because Tab Plendil ER is available in strength of 2.5mg [11]. Comparison of the therapeutic state transitions (both for the practice overall and for individuals) with respect to the selected guidelines was used to formulate criteria for a quality audit report.

Results

A State-Transition Overview Diagram (STOD) for the investigated period was computed based on all AHT state variable changes. Figure 1 shows all transitions that occurred at least 5 times. *Init_out* denotes the state when the patient has had no antihypertensive prescriptions for a minimum of 100 days before the start of AHT during the investigated period, otherwise the patient commences with *Init_in*. It is hypothesized that Init_out patients are more likely to be commencing AHT, whereas Init_in are more likely to be already in AHT at the time of their first prescription in the data extract. A "lapse" in antihypertensive therapy implies the period which commences when all antihypertensive medications, if taken as directed from the day of prescribing, should have run out, and is indicated by the *Zero* state. States are heuristically processed to avoid over-sensitivity [5], with Zero states of less than 90 days being coalesced into the prior state, as is any other state of less than 30 days duration.

Individual Path Diagrams (IPDs) were computed to illustrate the therapeutic experiences of individual patients as identified under a variety of exploratory criteria. Criteria focused on: (a) cycles (returning to the same state one or more times) and (b) transitions that are difficult to align with guidelines. With respect to the latter case, effective combinations for AHT include ACEi/ARB agents with diuretic, beta-blocker with diuretic and beta-blocker with DCCB [8]. Paths that break effective combinations may possibly adhere to best-practice, but are considered worthy of further scrutiny. A selection of interesting IPDs from the data set is given in Figure 2. Graphviz (URL: http://www.graphviz.org/) was used to produce Figures 1 and 2.

The STOD and individual IPDs were analyzed to identify relevant statistics for an audit report, both supportive and cautionary with respect to AHT quality. States and transitions were compared to the guidelines (note both RG and TK are physicians) as a means of knowledge engineering a TAR that aligns with the selected guidelines. In the present study, this was an exploratory exercise drawing firstly from the STOD and subsequently by assessing specific IPDs to confirm appropriate focus areas for quality improvement. For example, a transition of Init_in A for patients with diabetes as well as hypertension is readily aligned as an attribute of best-practice. Presence in the B1 state (or any combination state that includes B1) for

Table 2 - Classification of Therapeutic state variables and their respective Anatomical Therapeutic Chemical (ATC) classifications

AHT state variables	AHT drug groups (drug names for matching to PMS data)	ATC codes (from [12])
A	Angiotensin converting enzyme inhibitors(ACEi) and Angiotensin receptor blockers(ARB) (Enalapril, Candesartan cilexetil, Losartan potassium, Cilazap, Lisinopril, Quinapril, Accupril, Trandolapril, Perindopril erbumine, Captopril, Inhibace)	ACE inhibitors, plain: C09A and Angiotensin II antagonists, plain C09CA
B1	Beta-blockers (Atenolol, Esmolol hydrochloride, Acebutolol hydrochloride, Nadolol, Timolol maleate*, Propranolol, Carvedilol, Labetalol hydrochloride, Celiprolol hydrochloride, Betaloc, Metoprol, Oxprenolol hydrochloride, Pindolol)	Beta blocking agents, non-selective: C07AA and Beta blocking agents, selective: C07AB
B2	Diuretics (Bendrofluazide, Hydrochlorothiazide, Amiloride, Frusemide, Triamterene, Spironolactone, Bumetanide, Indapamide hemihydrate)	Low-Ceiling Diuretics, Thiazides: C03A, Low-Ceiling Diuretics, Excl. Thiazides: C03B, High-Ceiling Diuretics: C03C, Potassium-Sparing Agents: C03D, Diuretics And Potassium-Sparing Agents In Combination: C03E
B3	Non-dihydropyridine calcium channel blockers (Non-DCCB) (Verapamil, Dilatiazem)	Calcium-channel blockers: Selective Calcium Channel Blockers With Direct Cardiac Effects: C08D
C	Dihydropyridine calcium channel blockers (DCCB) (Felodipine, Isradipine, Nifedipine, Amlodipine)	Calcium-channel blockers: Dihydropyridine derivatives: C08CA
D	alpha blockers, hydralazines, Clonidine (Prazosin hydrochloride, Terazosin hydrochloride, Doxazosin mesylate, Clonidine, Hydralazine hydrochloride, Diazoxide)	alpha-adrenoreceptor antagonists: C02CA, Hydralazine: C02DB02, Clonidine: C02AC01

* prescribed as an Antihypertensive agen

patients with asthma is contraindicated and can contribute to a cautionary statistic.

The alignment of specific therapeutic states and transitions to the selected guidelines and best-practice were discussed within the author team. At times all or a subset of patients in a state- transition-based cohort were rematched by TK and reviewed by staff of the practice. The states and transitions were found to act as 'frames' for relatively simple sets of additional qualifiers. That is, almost no state or transition is simply good or bad; however, many clearly fit a supportive or cautionary category after the identification of a qualifier in terms of comorbidity or observation values.

Quality indicators (QIs) identified for a TAR fell into three broad categories:

1. Guideline based – QIs stemming from outcome or process requirements with limited relationship to states and transitions (e.g., achieving target BP).
2. State-transition based – QIs based more-or-less entirely on states and transitions with limited reference to other PMS data; notably, breaking of effective combination therapies.
3. Hybrid – a large set of QIs where state or transition is qualified by comorbidity or observations; these subdivide into 'bad' states (e.g., presence in Zero state without support of acceptable BP observations) and problem-drug interactions (e.g., asthma with beta-blocker).

The resulting report based on the TAR model has the following sections which combine the three types of QIs above:

- I. Description of practice as per PMS data (15 QIs):
  - General (e.g., patient volume)
  - Hypertension (e.g., prevalence)
  - Antihypertensive Therapy and Monitoring
- II. Criteria to support AHT quality (15 QIs):
  - Blood pressure control in patients classified with Hypertension
  - Continuity of therapy in patients classified with Hypertension

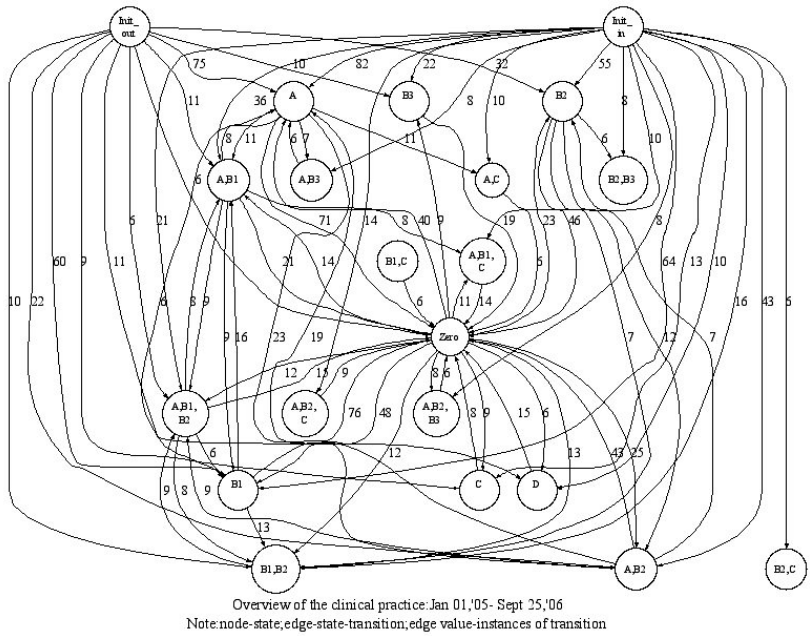


Figure 1 - State-Transition Overview Diagram (STOD) of antihypertensive therapy from 1 Jan 05 to 25 Sep 06, showing therapeutic states (circles) and transitions (arcs) with number of transitions made by patients during the investigated period.

  - Effective combination therapy in patients classified with Hypertension
  - Drug-problem indication in patients classified with Hypertension
  - Monitoring
- III. Criteria for improvement/caution of AHT (20 QIs):
  - Blood pressure management and monitoring
  - Lapses in antihypertensive therapy
  - Drug-drug interaction (e.g., concurrent beta-blocker and Diltiazem with or without atrial fibrillation and/or flutter)
  - Drug-problem interaction

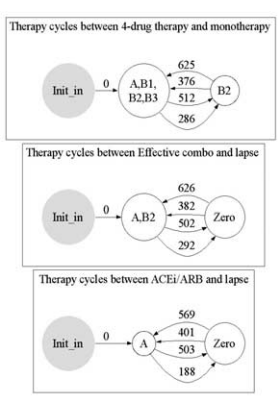


Figure 2 - Individual Path Diagrams (IPDs) showing cyclical patterns in antihypertensive therapy for select patients (circles indicate states; arcs labelled with day of transition relative to start of therapy).

Discussion

The TAR and associated analysis process supports a variety of quality assurance related activities. Broadly, these concern: (a) improvement of care for specific patients in the near term; (b) reflection on guidelines, practice and alerts; and (c) the use of QIs in relation to benchmarking and performance incentives.

Direct use of the TAR is for follow-up on the cohorts of patients responsible for critical indications. Actions may include review of the PMS record, and possible patient recall or review of therapy at next visit. The practice with which we have participated in this study is currently in the process of utilizing our report in this fashion.

A second use of the TAR is to support practice introspection beyond the immediate needs of specific patients. TAR QIs are tightly linked to concordance with guidelines. The TAR formulation process encourages clinicians to put the individual clauses of a guideline into context and to specifically qualify and interpret each clause for their practice. The extent of non-concordance to the guideline interpretation is provided in the report, and thus the potential value and impact of employing strategies such as local information programmes or online alerts can be assessed.

The tertiary use of the TAR is for benchmarking – that is, to compare the QIs to those for other practices. QIs emerging from the TAR model have several good characteristics in that they are tightly bound to clinical practice guidelines, have precise and compact definitions, and can be computed from PMS data without further clinician effort.

Relatively low performance against comparable clinics on such QIs should prompt consideration for quality improvement efforts.

The most controversial use of QIs (not particularly those from the TAR model) is as a basis for pay-for-performance funding incentives. Such incentives are already a reality in the UK [13] and substantial portions of the US [14] primary healthcare systems. Moreover, Teich et al. [15] suggest funding and incentives to reward the use of advanced decision support features in electronic prescribing, such as online alerts. Use of the TAR model may help to inform decision makers on the universe of possible QIs to achieve those with strong alignment to evidence-based best practice.

The method presented herein has a number of limitations. A number of manual steps were involved; however, we believe these could be minimized with replication, especially for key domains such as AHT. The method as applied in this study is limited by the quality of coding practice – our current results are based on the (questionable) assumption that all relevant conditions, comorbidities and observations can be found in the PMS record. Moreover, of course, the method is only applicable in an environment that has adopted General Practice computing.

Finally, it should be acknowledged that therapeutic state-transition based analysis covers some of the same ground as any other approach to guideline engineering and, conversely, that not all aspects of quality-of-care are based on therapy in the narrow sense we have applied it (e.g., there is quality of monitoring, and achieving patient ownership of care plans).

Conclusions

The TAR model report formulation process can be utilized to develop quality indicators (QIs) for audit of AHT in a clinical practice on the basis of changes in a patient's therapy over time as indicated in electronic records resulting from routine care activities. Uses can be characterized as including immediate review of therapy for specific patient cohorts, broader-view refinement and introspection of the uptake of guidelines into a practice, benchmarking to other practices, and strategic considerations on which QIs may be the best basis for incentives to promote evidence based care.

The direction of the current research is to look at areas for roll-out of the current TAR to a spectrum of practices; in New Zealand, this is best pursued through PHOs (Primary Health Organisations). We are also examining applicability to other CVD risk factors, notably dyslipidaemia agents.

References

[1] Didham R, Martin I, Wood R, and Harrison K. Information Technology systems in general practice medicine in New Zealand. N Z Med J. 2004 Jul 23;117(1198):U977.

[2] Kuperman GJ, Teich JM, Gandhi TK, and Bates DW. Patient safety and computerized medication ordering at Brigham and Women's Hospital. Jt Comm J Qual Improv. 2001 Oct;27(10):509-21.

[3] Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, and Strom BL. Role of computerized physician order entry systems in facilitating medication errors. *JAMA*. 2005 Mar 9;293(10):1197-203.

[4] Turley M, Stefanogiannis N, Tobias M, van der Hoorn S, Lawes C, Ni Mhurchu C, and Rodgers A. Nutrition and the burden of disease: New Zealand 1997-2011. Wellington, Ministry of Health, August 2003. Public Health Intelligence Occasional Bulletin Number 17.

[5] Warren J, Gadzhanova S, Stanek J, and Iankov I. Understanding caseload and practice through analysis of therapeutic state transitions. AMIA Annu Symp Proc. 2005;:784-8.

[6] Gadzhanova S, Iankov I, Warren J, Stanek J, Misan, G, Baig Z, and Ponte L. Developing High-Specificity Anti-hypertensive Alerts by Therapeutic State Analysis of Electronic Prescribing Records. J Am Med Inform Assoc. 2006 Oct 26; [Epub ahead of print]

[7] Texas Department of State Health Services: Hypertension Algorithm for Diabetes Mellitus in Adults, Revised 1-26-06. URL: http://www.dshs .state.tx.us/diabetes/PDF/algorithms/HYPER. PDF. Retrieved July, 2006.

[8] Therapeutic Guidelines Ltd., Australia. Therapeutic Guidelines: Cardiovascular. Version 4. 2003.

[9] Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL Jr, Jones DW, Materson BJ, Oparil S, Wright JT Jr, and Roccella EJ. Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. Hypertension. 2003 Dec;42(6):1206-52.

[10] The National Heart Foundation of Australia: Hypertension Management Guide for Doctors 2004. URL: http:// www.heartfoundation.com.au/downloads/ hypertension_management_guide_2004.pdf. Retrieved July, 2006.

[11] MIMS Online. URL: http://www.library.auckland.ac.nz/ databases/learn_database/public.asp?record=Mims. Accessed Oct, 2006.

[12] WHO Collaborating Centre for Drug Statistics Methodology: ATC/DDD Index 2006. URL: http:// www.whocc.no/atcddd /indexdatabase/. Accessed June, 2006.

[13] Doran T, Fullwood C, Gravelle H, Reeves D, Kontopantelis E, Hiroeh U, Roland M. Pay-for-performance programs in family practices in the United Kingdom. N Engl J Med. 2006 Jul 27;355(4):375-84.

[14] Rosenthal MB, Landon BE, Normand SL, Frank RG, and Epstein AM. Pay for performance in commercial HMOs. N Engl J Med. 2006; 355(18):1895-902.

[15] Teich JM, Osheroff JA, Pifer EA, Sittig DF, Jenders RA; The CDS Expert Review Panel. Clinical decision support in electronic prescribing: recommendations and an action plan: Report of the joint clinical decision support workgroup. J Am Med Inform Assoc. 2005 Jul-Aug;12(4):365-76.

Address for Correspondence:

Prof. Jim Warren,
Chair in Health Informatics
Department of Computer Science – Tamaki
University of Auckland, Private Bag 92019,
Auckland, New Zealand
Tel: +64 9 373 7599/ Fax: +64 9 373 7503
Email: jim@cs.auckland.ac.nz

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Development of Case-based Medication Alerting and Recommender System: A New Approach to Prevention for Medication Error

Kengo Miyo[a], Yuki S. Nittami[a], Yoichiro Kitagawa[a], Kazuhiko Ohe[a]

[a] *Department of Planning, Information and Management, the University of Tokyo Hospital, Japan*

Abstract

The purpose of this study was to develop a new alerting and recommender system for preventing medication errors. In recent years, alerting systems have been widely implemented, but because these systems apply a same static threshold for all patients in all cases, they produce excessive alerts and subject physicians to "alert fatigue". We believe that the most commonly-written prescription for a patient's status is the safest one. From this standpoint, we developed a real-time case-based medication alerting and recommender system linked to a database of past prescriptions. When a physician issues his or her prescription, our system dynamically compares it with past ones for similar patients in the database. An analysis of the 10 most frequently-used drugs in the University of Tokyo Hospital revealed that our system reduced the number of false alerts compared to the traditional static alert method. Our system contributes to the creation of alerts that are appropriate for patients' clinical conditions and based on physicians' empirical discretion.

Keywords:

case-based alerting; decision support systems, clinical; medical order entry systems; medication errors; prescriptions, drug

Background

In recent years, computerized physician order entry (CPOE) systems have been introduced to health care institutions worldwide [1, 2]. In Japan, the use of CPOE systems for medication has become widespread [3], and their implementation rate in hospitals with 500 or more beds was 70% as of 2005 [4]. The use of CPOE system is expected to contribute to efficient health care delivery and reduce physicians' time costs [1, 3, 5].

CPOE systems also contribute to improvements in health care quality, particularly in regard to safety [6–8]. To decrease errors in medical treatments, several systems with real-time data input checks have been developed, for example, to detect inappropriate dosages or drug combinations. In these systems, as soon as a physician clicks the "issue prescription" button, the system compares input data on dose regimen and concomitant drugs to data on dosage limitations and contraindicated drugs stored in a master table file. If the prescription contains inappropriate data, the system displays an alert. Alert systems that use static threshold data stored in a master table file are called *static alert systems*.

Although static alert systems are generally useful, excessive alerts can be produced because the systems check off patient status and treatment policies. Excessive alerts cause physicians to pay less attention to the alerts [9], which then lose their effectiveness. In other words, the few important alerts are overlooked amid a lot of meaningless ones. Peterson et al. described physicians in this situation as being in a state of "alert fatigue" [7, 10]. Particularly in university hospitals, where doctors treat many patients whose cases run counter to standard treatment, the risk of overlooking alerts cannot be ignored.

Physicians need to receive appropriate alerts. In this paper, the term *appropriate alert* refers not to an alert generated based on whether the data match standards determined by drug notes, but rather to an alert generated based on whether the present treatment differs greatly from actual treatment records. Our purpose in this paper is to suggest a new approach for generating appropriate prescription alerts. We believe that most commonly written prescription is the safest one, and we considered past records stored in HIS as the gold standard. From this standpoint, we developed a real-time case-based system that alerts physicians when their prescription deviates from this gold standard.

System design

System structure

The overview of our system is shown in Figure 1. The system consists of existing HIS and a database of past prescription records, as well as an alert engine. MySQL version 5.0 was chosen as the database for our system. All prescription data from the University of Tokyo Hospital (UTH) for 2000 through 2005 were extracted from HIS and stored in the database. For the alert engine, both Perl version 5.8, PHP version 5.1, and Apache version 2.0 were used.

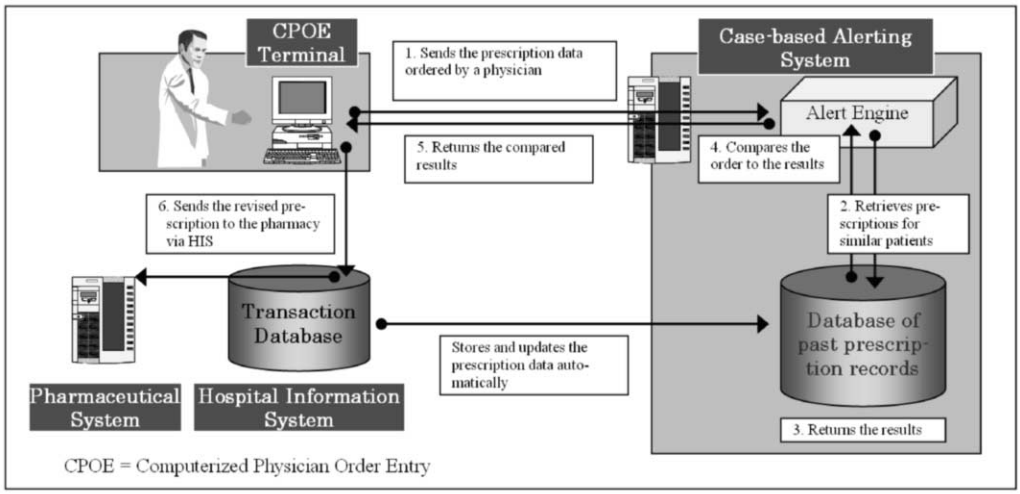


Figure 1 - The structure and workflow of the case-based alert system.

System feature

Our system has two major functions: alert function and decision support.

Alert function

The alert engine receives the prescription data entered by a physician before sending them to the pharmaceutical system. The alert engine compares these data to the statistical data for that same drug, compiled from past records in the database. If the alert engine finds that the entered data exceed the threshold described below, it sends an on-screen alert to the physician. Figure 2 presents an example of such a prescription alert. At present, the items that can trigger an alert are dosage, dose regimen, duration of drug administration, and concomitant drugs.

Statistical data on past prescriptions (such as 95th percentile and standard deviation) serve as the alert threshold. Prescription data in the database are also linked to the patient in question and to his or her diseases. Therefore, the alert thresholds classified according to clinical department (specialty), disease, sex, and age can be also used.

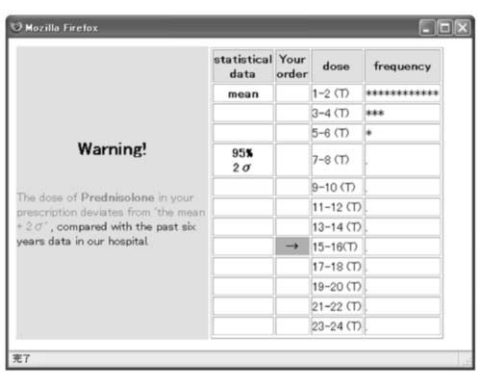


Figure 2 - A pop-up alert on the computerized physician's order entry terminal.

Decision support function

The recommendation function shown in Figure 3 is used when a physician wants to know the appropriate dosage and duration of drug administration. This function provides the most frequently used value by analyzing a distribution of values of past records. The physician can reexamine his or her prescription with these recommendations in mind. As in the case of the alert function, the recommendation function takes-the patient's attributes into consideration.

Figure 3 - The window of the recommendation function

Interim system evaluation

The final evaluation of our system should be done by physicians who actually use it. The present system, however, is still in the prototype phase. Therefore, we here clarify some interesting aspects of our method by comparing it to traditional static alert systems. We used data on dosage for the evaluation.

Evaluation method

We used the September 2006 the UTH prescription records for the system evaluation. We compared cases using our method to those using information about dosage limitation provided by drug notes. The alert function of our system used two thresholds. One was the 95th percentile threshold

of each drug, and the other was the mean + 2. We analyzed the 10 most frequently prescribed drugs at the UTH.

Evaluation results

Results are shown in Table 1. Under both the mean + 2 and the 95th percentile thresholds, we found that the number of alerts declined for nearly one-third of the drugs compared to the existing static alert method. However, because the value of the mean + 2 for prednisolone is lower than the value described in the drug note, the number of alerts increased. The reason is that the range of prescribed dosages of prednisolone is broad (i.e., the maximum prescribed dosage was up to 120 mg, whereas the values of prescribed dosages were concentrated around 5 and 10 mg). Thus, our research showed that our system, based on physicians' empirical discretion, set different thresholds than the existing static alert methods.

Discussion

System evaluation

One advantage of our method is that the number of alerts distinctly decreased from that produced by traditional static alert methods because unlike these traditional methods, our thresholds have upper values (e.g., for drugs with high variance of dosage). Loxoprofen sodium and amlodipine are good examples. In the case of prednisolone and aspirin, the mean + 2 thresholds were lower than those described in the drug notes, which caused the increase in alerts. It is critical to warn physicians when they are about to prescribe drug dosages that they do not often prescribe, even if that dosage is within the normal range described in a drug note. This is another advantage of our case-based method. It is one of our key future tasks to determine which threshold to use: mean + 2 or 95th percentile.

Another advantage of our method is that it can issue alerts differently for each clinical department. Prednisolone, whose dosage varies greatly among departments, is a good example. Applying a threshold across the board as traditional methods do results in the production of unnecessary alerts in certain departments. Our system, however, avoids this issue by issuing alerts based on department features, and on values such as disease, sex, and age, even for one drug. This ensures that alerts are appropriate for each patient's clinical conditions.

Although our system is in the prototype stage at present, we will start actual operation in 2007 and survey physicians' prescribing behavior and degree of satisfaction.

Availability of case-based approach

Master table maintenance is unnecessary

The traditional method, which keeps static descriptions of dosage limit and contraindicated drugs in master tables, requires a huge amount of work to maintain these tables. Our system, however, requires less maintenance because information on dosage limit and contraindicated drugs is dynamically created using already stored data. In addition, our method keeps up with medical advances. When dose regimens change, our method automatically updates the alert thresholds and consequently can issue alerts based on the current prescription trends.

Table 1 - Results of Interim Evaluation

Drug	Total number of orders	Threshold from drug note	# of alerts	*M* + 2*	# of alerts	95th percentile*	# of alerts
Teprenone	3424	150 mg	14	202.3 mg	0	150 mg	14
Loxoprofen Sodium	2623	180 mg	27	244.8 mg	5	240 mg	5
Brotizolam	2666	0.25 mg	460	0.485 mg	447	0.5 mg	1
Famotidine	2238	40 mg	3	51.2 mg	3	40 mg	3
Prednisolone	2110	60 mg	32	41.9 mg	77	40 mg	77
Rebamipide	2094	300 mg	12	398 mg	12	300 mg	12
Amlodipine	1915	5 mg	392	10.2 mg	2	10 mg	2
Aspirin	1950	300 mg	0	125 mg	24	100 mg	24
Sennoside AB	1820	48 mg	26	47.2 mg	154	48 mg	26
Mecobalamin	1849	1500 µg	1	1815 µg	1	1500 µg	1

* The system issues an alert when the dosage is greater than this value.

Oriented to physicians' experiences

Our system promises to clarify various dose regimens in a way that is consistent with physicians' empirical discretion. By using past records, our system can alert physicians to drug combinations not empirically prescribed, even if the drugs are not described in drug notes as being contraindicated. In addition, our system can present information about recommended concomitant drugs, similar to the way interns receive advice from supervisors.

Possibility of graded alert

Our method provides distributions of various parameters such as dosage and dose period. As a result, our system can present graded alerts, such as 1: attention, 2: warning, and 3: prohibited. Gradings of how extensively prescription data diverge from past records enable the physician to quickly decide whether he or she should heed the message.

Limitations of our system

One limitation of our system is that it does not function well with an insufficient amount of past data. For drugs on which a hospital has few or no past records, it is necessary to combine our method with other existing techniques. Small hospitals that do not have a sufficient number of past records must import prescription data from authorized large-scale hospitals. Using data from other hospitals is equivalent to asking "In your hospital, how is this drug commonly used?"

Future direction

The alert and recommendation function can also be applied to laboratory test and radiological records. We are now considering how to make the system available for other medical practices.

Conclusions

We developed a real-time case-based medication alert system that alerts physicians when a prescription deviates from similar commonly written prescriptions in a database of stored records. We set past records stored in HIS as the gold standard. An analysis of the 10 most frequently prescribed drugs at the UTH revealed that our system reduced the number of unnecessary alerts compared to the traditional static alert method. Our easy-to-maintain system creates alerts appropriate for patients' clinical conditions and alerts based on physicians' empirical discretion. In addition, it can provide more appropriate alerts than traditional methods, and as a result, will contribute to the prevention of errors in medical treatments.

Acknowledgments

This study was supported in part by a Grant-in-Aid to Young Scientists (B) 18790353 (2006) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

[1] Kuperman GJ, and Gibson RF. Computer physician order entry: benefits, costs, and issues. Ann Intern Med 2003: 139(1): 31-9.

[2] Rothschild J. Computerized physician order entry in the critical care and general inpatient setting: a narrative review. J Crit Care 2004: 19(4): 271-8.

[3] Haruki Y, Ogushi Y, Okada Y, Kimura M, Kumamoto I, and Sekita Y. Status and perspective of hospital information systems in Japan. Methods Inf Med 1999: 38(3): 200-6

[4] Ministry of Health, Labour and Welfare, Japan. Survey of medical institutions in 2005. Oct 1, 2005; Available from: http://wwwdbtk.mhlw.go.jp/toukei/data/160/2005/toukeihyou/0005605/t0123581/A0064_001.html. Accessed: Dec 4, 2006. (in Japanese)

[5] Koppel R, Metlay JP, Cohen A, Abaluck B, Localio; AR, Kimmel SE, and Strom BL. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005: 293: 1197-203.

[6] Judge J, Field TS, DeFlorio M, Laprino J, Auger J, Rochon P, Bates DW, and Gurwitz JH. Prescribers' responses to alerts during medication ordering in the long term care setting. J Am Med Inform Assoc 2006: 13(4): 385-90.

[7] van der Sijs H Aarts J, Vulto A, and Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc 2006: 13(2): 138-47

[8] Kaushal R, Shojania KG, Bates DW. Effects of computerized physician order entry and clinical decision support systems on medication safety. Arch Intern Med 2003: 163(12): 1409-16.

[9] Glassman PA, Simon B, Belperio P, and Lanto A. Improving recognition of drug interactions: Benefits and barriers to using automated drug alerts. Med Care 2002: 40(12): 1161-71.

[10] Peterson JF, and Bates DW. Preventable medication errors: identifying and eliminating serious drug interactions. J Am Pharm Assoc 2001: 41(2): 159-60.

Address for correspondence

Kengo Miyo, Ph.D., R.N., P.H.N.
Hongo 7-3-1, Bunkyo-ku,
Tokyo 113-8655, JAPAN;
Email: miyo-sup@h.u-tokyo.ac.jp>

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Is the Future Evidence-Based?

Ben W.I. Sissons[a], W. Alex Gray[a], Tony Bater[b], Dave Morrey[b]

[a] *School of Computer Science, Cardiff University, UK*
[b] *Clinical Information Unit, Velindre NHS Trust, UK*

Abstract

This paper is concerned with how the future information needs of the medical community should be met. The current dominant belief within medicine is that these information needs should be met from bespoke research studies. The necessity of this approach is far from certain. Health organisations worldwide are currently investing vast resources into centralising and amalgamating every day patient data. Is there a future for these Electronic Medical Records in informing medical decisions? This paper describes the challenges to be met in using both research studies and every day patient data to inform medical decisions. It then describes an ongoing practical project to evaluate these information sources' ability to meet the information needs of cancer care decision makers. Details of background, methodology and initial promising results are presented.

Keywords:

evidence-based medicine, clinical decision support systems, computerized medical records

Introduction

For healthcare providers worldwide the creation of national Electronic Medical Records (EMR) is seen as an integral part of meeting the challenges healthcare will face over the next 20 years. These record systems are being designed primarily to make administrative patient information available at all points of care. These systems will also develop into unparalleled sources of potential information for decision makers. This paper presents challenges and opportunities for the use of EMRs to inform decisions about patient treatment. The paper also highlights the need for re-assessment of current best decision making practice in light of these issues. It then describes a project to explore the role of EMRs in our healthcare future and its initial results.

The information needs of treatment decisions

The primary goal of informing treatment decisions is to support clinician and patient in identifying the most beneficial therapy for a patient. Throughout this paper the expression decision makers is intended to mean clinicians and patients.

Clinicians have a professional obligation to make all decisions in the best interests of the patient. Sentiments to this effect have been an integral part of medicine since the days of Hippocrates, circa 400BC. Making decisions primarily to suit the best interests of the patient was part of the original Hippocratic Oath and is a central tenet of the modern version, the declaration of Geneva.

In modern medicine it is necessary that any conclusions reached or advice given by health professionals be justifiable. Over the past few decades the professional obligations of clinicians to work in the best interests of patients has developed into a legal obligation. It is theorised that mass education, mass media and a changing legal culture have led to an increasing trend amongst patients to question the decisions of health professionals and seek legal redress for clinical negligence [1-3]. Evidence seems to support this, in the United Kingdom, for example, the number and cost of legal actions against the National Health Service has risen dramatically in the past decade [4, 5]. This increase in legal action, together with a number of high profile mistakes, has created a need for clinicians and healthcare providers to be able to demonstrate that decisions made or advice given to patients are based on relevant and valid information [6]. For ethical, moral and legal reasons evidence to support treatment decisions needs to be auditable, relevant and reflect how beneficial or harmful therapies are to patients.

Evidence-Based Medicine and the research study

Evidence-Based Medicine (EBM) arose to meet the information needs of medical decision making. It advocates basing decisions primarily on randomised and prospective data obtained from research studies. It is now the dominant methodology for informing medical decisions.

The EBM movement arose due to the need to better justify medical decisions. EBM was originally put forward as a paradigm shift from a model of medicine justified by practitioner experience to one based on the "best available" scientific evidence [1].

EBM advocates that treatment decisions should be based on analysis of specially designed studies. The proponents of EBM recommend the use of a strict hierarchy of evidence for informing medical decisions. For treatment decisions this hierarchy is topped by analysis of data from Randomised Controlled Trials (RCT). Where such analy-

ses are not available or cannot be used it is recommended that analyses of specially designed studies that collect data on "cohorts" of patients be used. Only in the absence of such research evidence should the analysis of retrospective, every-day patient data be considered as a basis for medical decisions [2-5]. EBM is very clear that evidence from research studies constitutes the best basis for making medical decisions.

EBM currently dominates the way in which medicine is practiced in much of the world.

Evidential support for the hierarchy of evidence

EBM's hierarchy of evidence is based on theories of good statistical practice; it is not backed by empirical evidence.

The modern hierarchy of evidence for treatment appears to have been proposed in a paper by several of the most prominent proponents of the EBM movement, *"Clinical Recommendations Using Levels Of Evidence For Antithrombotic Agents"* [3]. Preference is given to randomised experimental data because it is believed that without randomisation any conclusions about the relative benefits of treatments to patients would be clouded by the effects of unknown variables on outcomes [4, 5]. No primary evidence is provided for this theory in any of the above papers. This lack of evidence continues through the "users' guide to medical literature" series which represents the best practice manuals of EBM. However, all other factors being equal, it is logical to pay the most attention to the evidence which makes the greatest attempt to filter out bias from its results, in this case RCTs.

The use of analyses from prospective cohort studies is recommended where it is not feasible to use analyses of RCTs. The logic behind this is that, again all other factors being equal, a well planned study that collects its own data should yield more accurate results than a study making secondary use of data collected for other purposes. If all other factors were equal it would be difficult to disagree with this, a prospective study should be at least as sound as a retrospective one.

EBM's hierarchy of evidence seems therefore to be based on reasonable arguments but there is a lack of primary evidence to support these arguments.

Challenges of evidence-based decision making

It is difficult to assess the extent to which research evidence can be used to make inferences about patients in a separate, everyday clinical environment. This represents the greatest challenges in using research evidence to justify medical decisions.

The limited ability to generalise from research studies arises from the manner in which patients enter studies. Patients are recruited in medical centres that are asked to, and agree to, take part in the study. Eligible patients are then required to give their informed consent to take part in the study; in the case of RCTs this means doctor and patient must agree to let chance decide how the patient is treated. This selection process means the patients sampled for such studies are unlikely to be representative of the population they are drawn from, let alone a separate population of patients. Making inferences to decisions about patients in daily practice under such circumstances is poor statistical practice. Consequently the representativeness of medical studies is the major challenge faced by the Evidence-Based approach to informing medical decisions.

The problems with study recruitment and representativeness can be confirmed from practice. In the UK official figures show the participation of UK cancer patients in clinical trials is low and those running trials can find it difficult to recruit a full complement of patients [1]. A recently completed analysis of over 20,000 US cancer patients that looked at how representative participants in trials are of the population they are sampled from showed there is good reason for concern [2].

The assumption that it is possible to generalise conclusions from research studies, especially RCTs to patients in everyday practice appears to have little grounding in statistical practice or be supported by evidence.

The rise of the Electronic Medical Record

Healthcare organisations worldwide are investing vast resources into developing national EMRs. For example in England this takes the form of the National Program for Information Technology (NPFiT) [3], in Wales the Informing Healthcare project [4] and in Australia HealthConnect [5].

These national EMRs provide great opportunities for improving the information available to healthcare decision makers. The eventual goal of these projects is to make all relevant information about each patient available when and where it is needed. As a side effect massive standardised datasets covering the majority of encounters between patients and healthcare providers will be created. These massive datasets potentially represent all patients treated by each healthcare provider and will almost certainly contain information that can support medical decision making.

Challenges of Electronic-Medical-Record based decisions

Analysis of EMR data has the potential to be much more representative of local circumstance and patient outcomes than analysis from research studies. This data is retrospective and documents un-randomised treatment decisions. Consequently in an Evidence-Based Medical world EMR data is only fit for informing medical decisions in the absence of analyses from research studies. This section examines the challenges that would have to be addressed if EMR Data were to be used to inform treatment decisions.

Justifiability of results

It is very difficult to say from analysis of EMR data what causes any observed differences in patient outcomes. Hence it is difficult to compare how beneficial different treatments are from analysing such data. EMRs document everyday healthcare encounters. In most of these encounters decisions are unlikely to have been taken randomly. Consequently the decision about which treatment a patient receives will have been based on what is known about that

patient, both recorded and unrecorded data. It is difficult to know if any apparent benefit from a treatment is due to the treatment itself or the factors used to select patients for the treatment. For example if one treatment is given to physically fit patients and one to less fit patients, is it patient fitness or the treatment that causes those receiving the first treatment to have better outcomes?

Data quality

Data from EMRs are unlikely to be as suitable for a given analysis as data collected specifically for that analysis. Everyday clinical data is likely to be entered and amended in a high pressure environment where treatment of patients is a higher priority than following data entry protocols. Consequently the collected data may contain frequent errors and omissions. Additionally such data are, in general, recorded primarily for supporting daily care, not analysis. This means that any analysis is restricted to working with what has been recorded, not what may be the most relevant information. Therefore without great effort and foresight, such retrospective data is unlikely to be as comprehensive and suitable for analysis as a prospective dataset.

Representativeness

Although EMR data may document all appropriate past-patient cases, these are still not entirely representative of new patients. Any inferences from the data must still be made across time. It can be argued consequently that EMR data is potentially a more representative base for real-world treatment decisions, but it is still far from perfect in this regard.

Summary

With the rising interest in National EMRs come new opportunities to better inform medical decisions. Unresolved and unmeasured problems exist in adopting either an Evidence-Based or an EMR-Based approach to informing medical treatment decisions. The need for justifiability in medical decisions must be born in mind. Currently an Evidence-Based approach can be justified by its acceptability within medicine. An EMR-Based approach must go to greater lengths to justify its conclusions. It is therefore necessary to develop an EMR-Based approach to informing medical decisions which is capable of addressing the challenges identified in the previous section. Any approach developed then needs to be evaluated against alternative methods for informing medical decisions. The remainder of this paper is concerned with a project that attempts to achieve these objectives.

Methods

A description of an ongoing project in South East Wales, UK, to assess the future role of EMR data in informing cancer care decisions is presented.

Objectives

The broad objectives of this project are:

- To analyse how data from EMRs can justifiably inform medical treatment decisions.
- To compare the practical usefulness of decision information models based on EMRs with models based on best available research evidence.

Methods

Specifically the project examines the role EMRs and best available research evidence could have played in informing past cancer care decisions in South East Wales, UK.

A series of decision information models based on analysis of EMRs and best available research evidence will be used to inform simulated decisions about cancer patients whose outcomes are already known.

Each model will be evaluated against three overall criteria:

1. **Information Provided** - What information would the model have provided to decision makers?
2. **Justifiability** - How justifiable a basis for decisions would the model have been?
3. **Accuracy** - How accurate would the outcomes information provided have been?

Evaluation

As stated above, models will be built and evaluated for information provided, justifiability and accuracy of predictions. This section describes the planned evaluation of each of these factors.

Information provided

This evaluation is designed to determine what benefits or drawbacks each information source has for decision makers. Each of the models will be examined to see what types of information they provide or fail to provide, what form this information is provided in and how specific the information is. This evaluation will be carried out on a qualitative basis with the help of clinical professionals and patient representatives.

Justifiability

This evaluation is to determine to what extent models provide an auditable and defensible basis for decision making. Models will be evaluated for the clarity of conclusions reached and the guarantees or measures of error they offer. Particularly important considerations will be how effectively each model addresses criticisms levelled at it and if the assumptions made in employing the model are explicit and testable. This evaluation will also be qualitative, weighing the pros and cons of basing decisions on each model. Much of this discussion has already been covered above.

Accuracy

This is intended to be a quantitative evaluation of the ability of each model to correctly predict the outcomes of previously unseen patients. The evaluation will assess, with the benefit of hindsight, the practical usefulness of these models to real world decision makers.

In order to evaluate each model's predictive accuracy, a decision simulation will be carried out using historical data of cancer patients in South East Wales. Each model will be

used to make predictions about the outcomes of patients diagnosed in 2003 and 2004 over the two year period following the diagnosis. The average predictive error of the models will be compared to a control model to assess which would have been the most useful for predicting patient outcomes.

Materials

Datasets

Healthcare providers of the South East Wales Cancer Network have operated a shared EMR system for over a decade. The system is appropriately named Information System for Clinical Organisations (ISCO). The purpose of this system has been to collect high quality data for analysis as well as administrative purposes.

The study will be carried out on three data sets from ISCO; Colorectal cancer patients, Breast cancer patients and Oesophageal cancer patients. These datasets have been chosen for the availability and quality of data and research evidence

Models

For each type of cancer, a set of decision information models will be built and evaluated. One model will be based on the best available research evidence as assessed by the Cochrane Collaboration [1] and the National Institute for Health and Clinical Excellence (NICE) [2], the UK's Evidence-Based Guideline body. A second model will be designed to draw as many justifiable conclusions as possible from data of patients from before 2003 contained in the ISCO EMR system. These models will be compared to two control models, also based on the pre 2003 patient data. One will be designed purely to achieve a good predictive accuracy in order to evaluate if useful information is lost by employing the justifiable past-patient data based model. The second will simply predict all outcomes as the average of the past patient dataset in order to establish if any of the other models can provide a benefit over this.

Measuring accuracy

The predictive accuracy of models will actually be measured as predictive error. Models will be tested against each other, making survival predictions for random samples of patients from the test dataset. The null hypothesis that there exists no difference in the underlying predictive error of each model will be assessed using matched pair t-tests. If there is strong evidence against the null hypothesis the alternate hypothesis that some models are better predictors than others will be accepted.

The measure of error will be Graf et al.'s Censored Brier Score [3]. This metric has been chosen as it is a measure of mean squared error and can cope with the special conditions of survival data. The objective of this evaluation is to measure the usefulness of decision information sources to those making decisions about individual patients. A measure of mean squared predictive error will be used as this takes greater account of accuracy for every individual case than non-squared error metrics. As the main outcome to be predicted is death from cancer, the metric must also be able to handle censored observations. These are cases in which the patient is still alive, out of contact or has died of other causes and so no outcome has been measured. Censoring rates can be over 50% of patients in cancer datasets. Ignoring censored cases is consequently likely to detract from the validity of the results. For a full description of the metric together with formulas see [3].

Initial results and discussion

This paper describes an ongoing investigation; the full investigation cannot be completed until 2007 as two years of outcomes data is needed for all test patients. Initial experimentation however has yielded some encouraging results.

Prototype systems based on the above models have been reviewed by clinicians to assess their ability to inform decision making. Feedback from this process has been good but a formal evaluation of finalised models has yet to be carried out.

The pros and cons of the justifiability of each data are discussed extensively above. Initial prototyping suggests an EMR-based system that addresses justifiability criticisms is feasible. Our current approach is to try and isolate effects of individual variables, especially the effects of treatment. The main challenge faced here is that resulting models tend to provide vague information, but progress is being made.

Promising results have been forthcoming from building the control model that attempts to get the best predictive accuracy from the pre 2003 EMR data (see above). The Cox Proportional Hazards model [1], a regression model specifically designed to cope with survival data, was compared to an average past-patient outcome model in 10-fold cross validation. The Cox model was fitted by systematically testing combinations of variables on the training set and choosing the simplest model that was the best fit. The cross validation was run 20 times, with each model using the same 9 folds of training data to make predictions for the test fold. For each of the three datasets a t statistic based on the difference between average Integrated Brier Scores was calculated. The resulting probability that the Cox model had the same predictive error as the average model was less than one in 4 billion chance ($p < 2.5*10^{-10}$) for each dataset. In all cases the Cox models produced a lower predictive error than the control models. The reduction in the predictive error was 3.2%, 5.6% and 3.4% respectively for the Colorectal, Oesophageal and Breast cancer patient datasets. These benefits from the fitted Cox model may seem small in terms of percentages, but these results show that there is an extremely high probability that the EMRs contain information that can be of benefit to decision makers.

Criticisms of approach

This project tests models against data from EMRs and consequently also faces the challenges described above. This does not detract from the fact however, that this is a very practical study, measuring information models against

real-world data. At the end of the day these models are designed to inform such real-world decisions.

Conclusion

For ethical, moral and legal reasons it is important to have reliable and justifiable sources of information for decisions over patient care. Currently Evidence-Based Medicine is the dominant methodology for meeting these needs and advocates basing decisions primarily on analysis of well designed, bespoke research studies. Many countries see national, and perhaps international, Electronic Medical Record systems as having a pivotal role in meeting future healthcare needs. This means that increasingly large collections of everyday clinical data are becoming available to decision makers. Should this data be used to inform medical decision or are we to write off a potentially valuable information resource as bad evidence? The study described here represents a step along the road to resolving this issue.

Although the project is still in its early stages, the initial results are very promising. There is almost certainly information contained in the South East Wales Cancer Networks Electronic Medical Records that could benefit decision makers; the major challenge is unlocking it.

Acknowledgments

We would like to thank clinicians and information staff at Velindre NHS Trust and the University Hospital of Wales for their valuable insight and support.

References

[1] Eysenbach G and Jadad A. Evidence-based patient choice and consumer health informatics in the Internet age. J Med Internet Res 2001: 3(2): p. e19.

[2] Risk A and Dzenowagis J. Review of Internet Health Information Quality Initiatives. J Med Internet Res 2001: 3(4): p. e28.

[3] Sastry S and Carroll P. Doctors, patients and the Internet: time to grasp the nettle. Clin Med 2002: 2(2): p. 131-133.

[4] NHS Litigation Authority. NHS Litigation Authority, 2006. http://www.nhsla.com/, Accessed: 28/09/2006

[5] National Audit Office. Handling clinical negligence claims in England. London: National Audit Office, 2001.

[6] Changing Attitudes and the Response to clinical claims and Complaints. Clin Risk 2001: 7(2): p. 124-125.

[7] Guyatt G. Evidence-Based Medicine - A new approach to teaching the practice of medicine. JAMA 1992: 268(17): p. 2420-2425.

[8] Levels of Evidence and Grades of Recommendation. Centre for Evidence-Based Medicine, 2005. WWW: http://www.cebm.net , Accessed: 15/11/2006

[9] Cook DJ, Guyatt GH, Laupacis A, Sackett DL, and Goldberg RJ. Clinical recommendations using levels of evidence for antithrombotic agents. Chest 1995: 108(4): p. S227-S230.

[10] Guyatt GH, Sackett DL, and Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 1993: 270(21): p. 2598-601.

[11] Guyatt GH, Sackett DL, and Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. JAMA 1994: 271(1): p. 59-63.

[12] Vale C, Stewart L, and Tierney J. Trends in UK cancer trials: results from the UK Coordinating Committee for Cancer Research National Register of Cancer Trials. Br J Cancer 2005: 92(5): p. 811-4.

[13] Elting LS, Cooksley C, Bekele BN, Frumovitz M, Avritscher EBC, Sun C, and Bodurka DC. Generalizability of cancer clinical trial results. Cancer 2006: 106(11): p. 2452-2458.

[14] NHS Connecting for Health. The National Programme for IT Implementation Guide. London: DoH, 2005.

[15] Welsh Assembly Government. Informing Healthcare: Transforming Healthcare Using Information and IT. Cardiff: Welsh Assembly Government, 2003.

[16] HealthConnect. HealthConnect: An overview. Canberra: HealthConnect, 2004.

[17] The Cochrane Collaboration: The reliable source of evidence in health care. The Cochrane Collaboration, 2006. http://www.cochrane.org/, Accessed: 20/10/2006

[18] National Institute for Health and Clinical Excellence. A Guide to NICE. London: NICE, 2005.

[19] Graf E, Schmoor C, Sauerbrei W, and Schumacher M. Assessment and comparison of prognostic classification schemes for survival data. Stat Med 1999: 18(17-18): p. 2529-45.

[20] Cox D. Regression Models and Life Tables. J R Stat Soc 1972: Series B(34): p. 187-220.

Address for correspondence

Ben Sissons
Email: B.W.I.Sissons@cs.cf.ac.uk
School of Computer Science, 5 the Parade, Cardiff, CF24 3AA, UK

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Comparing Decision Support Methodologies for Identifying Asthma Exacerbations

Judith W Dexheimer[a], Laura E Brown[a], Jeffrey Leegon[b], Dominik Aronsky[a,c]

[a] *Department of Biomedical Informatics, Vanderbilt University, Nashville, TN, USA*
[b] *Vanderbilt Center for Better Health, Vanderbilt University, Nashville, TN, USA*
[c] *Department of Emergency Medicine, Vanderbilt University, Nashville, TN, USA*

Abstract

Objective: To apply and compare common machine learning techniques with an expert-built Bayesian Network to determine eligibility for asthma guidelines in pediatric emergency department patients.

Population: All patients 2-18 years of age presenting to a pediatric emergency department during a 2-month study period.

Methods: We created an artificial neural network, a support vector machine, a Gaussian process, and a learned Bayesian network to compare each method's ability to detect patients eligible for asthma guidelines. Our outcome measures included the area under the receiver operating characteristic curves, sensitivity, specificity, predictive values, and likelihood ratios.

Results: The data were randomly split into a training set (n=3017) and test set (n=1006) for analysis. The systems performed equally well. The area under the receiver operating characteristic curve was 0.959 for the expert-built Bayesian network, 0.962 for the automatically constructed Bayesian network, 0.956 for the Gaussian Process, and 0.937 for the artificial neural network.

Discussion: All four evaluated machine learning methods achieved high accuracy. The automatically created Bayesian network performed similarly to the expert-built network. These methods could be applied to create a real-time detection system for identifying asthma patients.

Keywords:

asthma, decision support techniques, decision support systems, evaluation

Introduction

Asthma is one of the most common pediatric illnesses. It causes an estimated 14 million missed school days and 14.5 million missed workdays yearly in the United States [1]. Asthma exacerbations account for more than 1.8 million emergency department (ED) visits annually [1].

Adherence with guidelines has been shown to improve the clinical care patients receive [2, 3]. A common barrier to guideline initiation is determining eligible patients [4]. Nurses in addition to their normal workload are often charged with identifying eligible patients, a task that is frequently forgotten and leads to lower guideline adherence. Automatically detecting patients could help improve guideline adherence. Ideally, an electronic system that requires no additional data entry would be used to detect guideline eligible patients.

The goal of our study was to evaluate several machine learning techniques using a verified asthma dataset and to compare the techniques with an existing expert-built Bayesian Network (BN) used to identify asthma patients with asthma exacerbations in a pediatric ED [5]. We compared the expert-built BN with a BN automatically learned from data, a Support Vector Machine (SVM), an Artificial Neural Network (ANN), and a Gaussian Process (GP).

Background and methods

Various computerized methods have been developed for identifying asthma patients [6, 7]. These studies have focused on clinical data - such as peak flow, clinic notes or discharge summaries, and computer-based questionnaires [6]. Although some of the studies used ED data, they were not stand-alone systems or integrated into the clinical workflow or with existing information technology infrastructures.

Setting

The study data were collected at Vanderbilt Children's Hospital ED, a 29-bed pediatric facility with more than 40,000 visits annually. The ED information system infrastructure includes an electronic whiteboard, an electronic triage system, an electronic medical record, and a computerized provider order-entry system in the ED. The study was approved by the institution's IRB.

Population

All patients presenting to the pediatric ED during the 2-month study period were screened for inclusion. Patients were included if they were 2-18 years of age and seen in the pediatric ED. Patients were excluded if they did not have a coded chief-complaint, were not treated in the ED (such as left without being examined) or had no final diagnosis for their visit in the paper or electronic patient record.

Design objectives

Our objective was to apply and compare machine learning techniques to an expert-built Bayesian Network. The result of each technique could be developed into a real-time system for detecting asthma patients presenting to a pediatric ED. We adapted the design objectives of a prior BN study to fit our additional techniques [5].

Data sources

We used the database of the previous study [5] which included 4,023 patient encounters. The randomized splits for training and testing of the original BN were preserved. The data included commonly available variables from the electronic medical record, electronic triage, and electronic whiteboard such as 141 ICD-9 coded chief complaints, a past medical history of asthma, billing codes, and asthma medications including beta-agonists, steroids and others. Chief complaints and asthma medication were identified using searches of free-text. Text matching has been shown to be an accurate method to determine the past medical history of asthma [8]. The past medical history of asthma was determined by searching the billing records for prior asthma diagnosis codes (ICD-9 493.*). The patient's chief complaint, acuity level, age, respiratory rate, and oxygen saturation came from the electronic triage. All of these variables are regularly captured and stored in the hospital's information systems. All data elements were available through the computerized hospital information systems, and no additional data collection was necessary.

Dataset

The dataset collected consisted of 11 variables: 3 variables (age, oxygen saturation, and respiratory rate) have continuous values, 5 variables are ordinal (acuity, billing codes, and the 3 medication variables), 1 variable (chief complaint) is categorical, and 2 variables (history of asthma and prediction variable of asthma eligibility) are binary. The dataset was discretized following the values given in Table 1 for use in constructing the expert BN.

BN learning algorithms, including the one used in this study, generally require complete datasets, i.e., no missing values. The frequency of missing data in several variables (shown in table 1) limited the choice of methods to handle the missing values (e.g., removal of cases, removal of variables, imputation). The "non-asthma" column includes all patients who were not diagnosed with asthma (n=3,638) and the "asthma" column includes patients diagnosed with asthma (n=385). The missing elements were encoded as an additional value a variable may take. The dataset was randomly split into a training set (n=3017) and testing set (n=1006) with equal proportions of asthma cases.

The ANN, SVM, and GP methods may take continuous values, therefore an alternative dataset was constructed for their use from the original data (i.e. prior to discretization). For this dataset missing data for the continuous variables were imputed using a k-nearest neighbors imputation method. For the medication variables, a missing value was treated as having no history of that medication found in the medical record. Missing acuity values were treated as an additional "0" value. All ordinal and categorical variables were encoded using 1-of-m encoding. Finally, the dataset was scaled to [-1,1] for the SVMs and [0,1] for the ANN technique. The dataset was scaled using Gaussian normalization for the GP.

Table 1 – Missing values

Variable	Values	Non-asthma (%)	Asthma (%)
History of Asthma	Present, Absent	0.0	0.0
Billing Codes	Number: 0, 1, >1	0.0	0.0
Meds: -agonists	Number: 0, 1, >1	51.3	45.2
Meds: Steroids	Number: 0, 1, >1	51.3	45.2
Meds: Other	Number: 0, 1, >1	51.3	45.2
Chief Complaint	141 unique	0.0	0.0
Acuity Level	ESI level: 1-5	1.21	0.26
Age Category	2-3, 4-6, 7-11, 12-18	0.0	0.0
Respiratory Rate	<20, 20-24, 25-29, 30-34, 35-39, >40	4.15	3.38
Oxygen Saturation	<91, 91-92, 93-94, 95-96, 97-98, >98	3.13	0.26

Reference standard

The reference standard for an asthma diagnosis was a free-text diagnosis of "asthma exacerbation," "status asthmaticus," "wheezing," or "reactive airway disease" [5]. A board-certified internal medicine physician determined the asthma diagnosis through manual chart review. Electronic and paper charts were searched for a diagnosis for every ED visit during the study period. Patients without a discharge diagnosis were not included in the study.

Bayesian network

A BN is formalism consisting of a directed acyclic graph with nodes representing variables and a joint probability distribution over the variables. BNs can be created by hand using expert knowledge. The network parameters for the BN are set by the training set and predictions are made on the testing set. BNs are advantageous in that the prediction inference algorithms handle missing data which is prevalent in clinical systems, and they allow an investigator to choose an optimal detection threshold balancing sensitivity and specificity. BNs have been used for disease detection and diagnosis [9].

Max-Min Hill-Climbing

As an alternative to creating the structure of a BN by hand using expert knowledge, a BN was constructed automatically from the data using machine learning techniques. Many algorithms exist for learning BNs; the MMHC algorithm has been shown to outperform on average a number of other prototypical and state-of-the-art BN learning algorithms in an extensive empirical evaluation and was therefore chosen for use in the analysis [10]. MMHC learns the structure of a BN given a dataset, after which the

network parameters are then estimated directly from the data. The MMHC algorithm is available in the Causal Explorer library[1] [11]. The training dataset was used to learn the network structure and estimate the parameters. The Norsys Netica™ API probabilistic inference algorithm was used to predict asthma in the test dataset; the same that was used in the expert BN. This test dataset was used to calculate an AUC for the method.

Support Vector Machine

SVM's, when used in a binary classification problem such as in this study, construct a maximal margin separating hyperplane to discriminate between the two classes of data [12]. SVMs make use of a kernel function to map the input data to a new space typically of much higher dimensions, where an optimization procedure is run to find the linear separating hyperplane. Several different kernel functions are often considered for classification tasks; the full polynomial and radial-basis kernel functions (RBF) were both used in this task. The full polynomial kernel takes two parameters: the degree d of the function, and C, a cost parameter. The RBF kernel also takes two parameters: the cost parameter C, and which determines the width of the function. SVMs have been used to classify clinical data and in clinical data analysis [13, 14].

The SVM classifiers were implemented using LibSVM [15]. The choice of parameters for the classifiers was optimized with stratified 10-fold nested cross validation over the training dataset using empirical area under the receiver operating characteristic curve as a performance measure. The values for each parameter was selected from the following ranges: $C - \{10^{-6}, 10^{-4}, 10^{-2}, 1, 10^{2}, 10^{4}\}$, $d - \{1, 2, 3, 4, 6, 8\}$, and $ - \{10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1\}$. The best classifier was selected and trained using the entire training dataset, and evaluated on the test data set.

Artificial Neural Network

ANNs are modeled after the brain's interconnecting neurons [16]. They use a non-linear approach to create statistical models. ANNs can be used to discover patterns in a dataset, and have been applied to identify asthma patients using responses to a questionnaire [6]. "Learning" occurs through adjusting the connection weights between nodes, finding a result, and then re-adjusting the connection weights.

Our network was developed using the Netlab [16]. A nested 10-fold cross-validation was employed to select the network architecture. The networks were trained and tested using gradient descent with adaptive learning back-propagation with mean squared error as the fitness measure. The best network was selected and trained on the entire training set. An independent test set was then used to calculate a receiver operating characteristic curve (ROC) and asthma probabilities.

Gaussian Process

A GP applies Bayesian techniques to an ANN to create a probabilistic structure that can be used to calculate probabilities. Bayesian methods have been applied to ANNs by placing prior probabilities on the weights of the network. Using Bayesian methods with ANNs elevates the need for a monitoring data set and allows parameters to be determined on the network being trained [17]. GPs are an extension to BNs, but they place prior probabilities on the function.

We used the Gaussian Processes for Machine Learning (GPML) toolbox in Matlab [17, 18] to develop a GP for asthma prediction. The hyperparameters were optimized using the supplied function in the GPML toolbox. We applied a Laplace's Approximation for a binary Gaussian process and selected the commonly used squared exponential covariance function. The hyperparameters are related to the squared exponential covariance function to determine the amplitude and shift of the function in space. The hyperparameter lengths are associated with each variable and determine how much the variable is can vary in its dimension.

Analysis

Performance was evaluated using ROC curves [19]. The ROC curve measures overall test performance and is obtained by plotting sensitivity versus 1-specificity. The area under the ROC curve (AUC) was the primary outcome measure [20] for all techniques with the exception of the SVM. We determined standard operational characteristics for each method including sensitivity, specificity, predictive value, and likelihood. For probabilistic systems, sensitivity was varied from 80%, 85%, 90%, and 95% to determine standard operational characteristics. To compare the methods with the expert-built BN, sensitivity was fixed at 90% as in the previous paper [5].

Results

There were 4,115 patient visits during the study period. Of these, 92 visits were excluded, and 385 (9.6%) had a final diagnosis of asthma. The patient demographics were reported in the previous study. The Bayesian network structure displayed in the left hand side of figure 1 was constructed manually using expert knowledge. The right hand side of figure 1 displays the network structure learned from the data using MMHC.

1 Causal Explorer containing MMHC can be downloaded at http://www.dsl-lab.org.

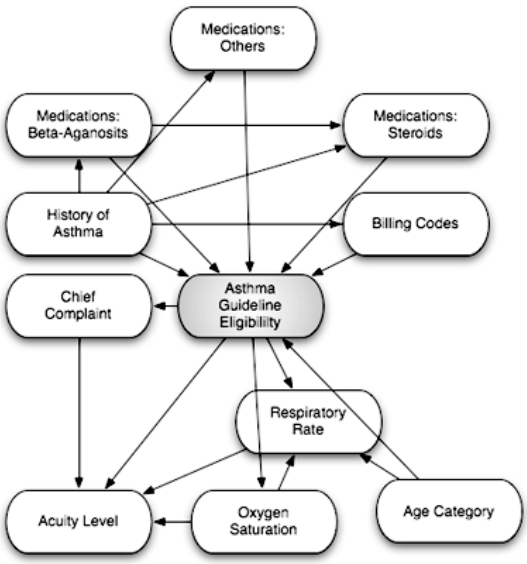
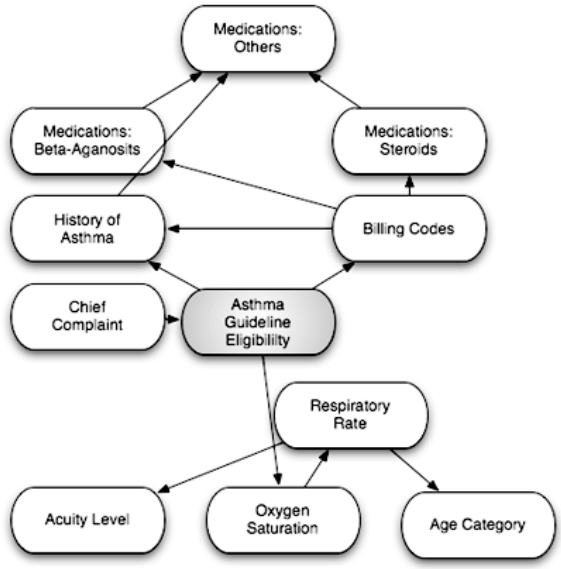


Figure 1 - Expert-built Bayesian Network



MMHC Bayesian Network (structure learned form data)

The AUC for the expert BN was 0.959 (95% CI: 0.933 – 0.977). The MMHC created network AUC was 0.962 (95% CI: 0.935 – 0.980). The ANN AUC with 160 hidden nodes and 160 hidden layers was 0.936 (95% CI: 0.902 – 0.961), and the AUC for the GP was 0.956 (95% CI: 0.923 – 0.976). The ROC curves for the original BN, the BN learned with MMHC, the ANN, and the GP are shown in figure 2. The SVM produces binary predictions with the threshold implicit in the SVM formalism therefore, an SVM ROC curve is not included in this portion of the analysis.

Table 2 shows operational characteristics of the methods with the fixed 90% sensitivity (for methods where the threshold can be varied). The results are reported directly on the predictions made by the SVM since sensitivity cannot be varied. The final parameters selected and used by the SVM was the full polynomial kernel function with a degree of 1 and cost parameter of 1. Operational characteristics for multiple sensitivities were calculated, as in the previous study, for the MMHC network are shown in table 3 for comparison.

Table 2 - Operational characteristics

	SEN (%)	SPEC (%)	PPV (%)	NPV (%)	PLR	NLR
BN	90.0	88.3	44.7	98.9	7.69	0.11
MMHC	90.0	90.1	49.2	98.9	9.16	0.10
ANN	90.0	84.4	38.0	98.8	5.81	0.11
GP	90.0	90.3	49.7	98.9	9.37	0.10
SVM	71.9	98.7	85.2	97.1	54.5	0.29

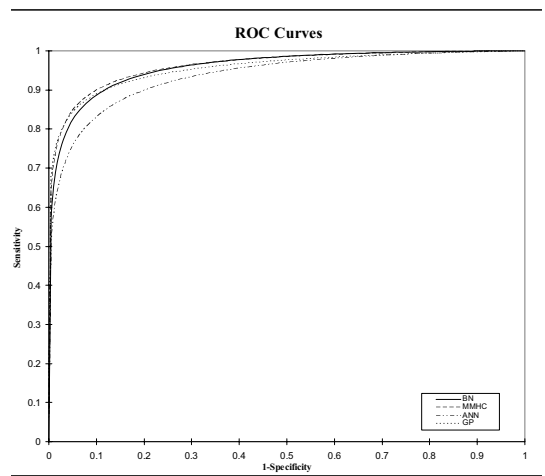


Figure 2 – ROC Curves

The test set had 96 asthma cases and 910 non-asthma cases. At 90% sensitivity, the expert-built BN had 86 true positive predictions and 105 false positive predictions, 805 true negative predictions and 10 false negative predictions. The MMHC network also had 87 true positive predictions and 90 false negative predictions, 820 true negative predictions and 9 false negative predictions. The ANN had 87 true positive predictions and 142 false positive predictions, 768 true negative predictions and 9 false negative predictions. The GP had 87 true positive predictions and 88 false positive predictions, 822 true negative predictions and 9 false negative predictions. At the resulting 72% sensitivity, the SVM had 69 true positive predictions and 12 false positive predictions, 898 true negative predictions and 27 false negative predictions.

Table 3 - MMHC operational characteristics with fixed sensitivity values

SEN (%)	SPEC (%)	PPV (%)	NPV (%)	PLR	NLR
80	97.1	74.8	97.9	28.1	0.20
85	93.8	59.1	98.3	13.7	0.17
90	90.1	49.2	98.9	9.16	0.10
95	86.3	42.1	99.4	6.90	0.06

SEN: Sensitivity, SPEC: Specificity, PPV: Positive Predictive Value, NPV: Negative Predictive Value, PLR: Positive Likelihood Ratio, NLR: Negative Likelihood Ratio

Discussion

The accuracy of the MMHC discovered BN, the SVM, and the GP were comparable to the expert-built BN, however, the ANN did not perform as well as the expert-built BN.

Sparse data may have caused problems in some of the techniques. In our dataset, asthma prevalence was 9.6% of the cases. With this few cases, the SVM and ANN, without adjustment, may not have been able to properly detect the asthma exacerbations. We did not perform any procedures for handling imbalanced data in this study (e.g., over-sampling, under-sampling, one class SVMs, etc). However, such adjustment may lead to better performance of the SVM and ANN.

The MMHC algorithm depends on tests of conditional independence. Accurate estimates for those tests depend on the number of samples and the domain of the variables involved in the testing. For this dataset the Chief Complaint variable can take one of 141 possible values. With the amount of sample provided to the learning algorithm caution must be taken when considering the network produced.

In summary, we believe these methods could be applied to create a real-time detection system for identifying asthma patients using commonly available clinical data.

Acknowledgments

The first author was supported by NLM T15 007450-03.

References

[1] Allergy & Asthma Advocate. Quarterly patient newsletter of the American Academy of Allergy, Asthma and immunology. 2004

[2] National Heart, Lung, and Blood Institute, National Asthma Education and Prevention Program. Expert Panel Report 2: Guidelines for the diagnosis and management of asthma. 1997

[3] Scribano PV, Lerer T, Kennedy D, Cloutier MM. Provider adherence to a clinical practice guideline for acute asthma in a pediatric emergency department. Acad Emerg Med. 2001;8:1147-1152.

[4] Grimshaw JM, Eccles MP, Walker AE, Thomas RE. Changing physicians' behavior: what works and thoughts on getting more things to work. J Contin Educ Health Prof. 2002;22:237-243.

[5] Sanders DL, Aronsky D. Detecting Asthma Exacerbations in a Pediatric Emergency Department Using a Bayesian Network. Proc AMIA Symp. 2006;:684-688.

[6] Sanders DL, Aronsky D. Biomedical informatics applications for asthma care: a systematic review. J Am Med Inform Assoc. 2006. [E-pub].

[7] Hirsch S, Shapiro JL, Turega MA, Frank TL, Niven RM, Frank PI. Using a Neural Network to Screen a Population for Asthma. Ann Epidemiol. 2001;11:369-376.

[8] Sanders DL, Gregg W, Aronsky D. Identifying asthma exacerbations in a pediatric emergency department: A feasibility study. Int J Med Inform. 2006. [E-pub].

[9] Aronsky D, Haug PJ. Diagnosing community-acquired pneumonia with a Bayesian network. Proc AMIA Symp. 1998;;632-636.

[10] Tsamardinos I, Brown LE, Aliferis CF. The max-min hill-climbing Bayesian network structure learning algorithm. Machine Learning. 2006;65:31-78.

[11] Aliferis CF, Tsamardinos I, Statnikov A, Brown LE. Causal Explorer: A Probabilistic Network Learning Toolkit for Biomedical Discovery. In International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS). 2003;:371-376.

[12] Vapnik V. Statistical Learning Theory. Wiley-Interscience; 1998

[13] Lagor C, Aronsky D, Fiszman M, Haug PJ. Automatic identification of patients eligible for a pneumonia guideline: comparing the diagnostic accuracy of two decision support models. Medinfo. 2001;10:493-497.

[14] Kote-Jarai Z, Matthews L, Osorio A, Shanley S, Giddings I, Moreews F, Locke I, Evans DG, Eccles D, Williams RD, Girolami M, Campbell C, Eeles R. Accurate prediction of BRCA1 and BRCA2 heterozygous genotype using expression profiling after induced DNA damage. Clin Cancer Res. 2006;12:3896-3901.

[15] Chang CC and Lin CJ. LIBSVM: a library for support vector machines. 2001

[16] Nabney IT. Netlab. Springer; 2004

[17] MacKay DJC. Information Theory, Inference & Learning Algorithms. Cambridge University Press; 2002.

[18] Matlab® v 7.3. http://www.mathworks.com

[19] Hanley JA and McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143:29-36.

[20] Hand DJ and Till RJ. A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems. Machine Learning. 2001;45:171-186.

Address for correspondence

Judith W Dexheimer, 2209 Garland Ave, Nashville, TN, 37232-8340, judith.dexheimer@vanderbilt.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Analysis and Redesign of a Knowledge Database for a Drug-Drug Interactions Alert System

Daniel Luna, Victoria Otero, Daniela Canosa, Sergio Montenegro, Paula Otero, Fernan Gonzalez Bernaldo de Quirós

Department of Medical Informatics, Hospital Italiano de Buenos Aires, Argentina

Abstract

Physicians tend to ignore drug-drug interactions alerts, this is due to the large amount of irrelevant interactions displayed and the interface in which these alerts are shown. The high rate of clinically inadequate alerts produce "alerts fatigue". This high number of incorrect alerts predisposes physicians to underestimate the electronic prescription systems as useful tools in their practice. We decided to analyze and redesign our drug-drug interactions alerting system knowledge database. In order to do so, we cleaned our knowledge database according to the clinical significance of drug-drug interactions. New drug interactions taxonomy was created in four levels based on clinical significance and the recommendations given in each single monograph of interaction. We proceeded to recategorize the alerts as Active, which present themselves to the physician interrupting the prescribing process, or Passive, which allow physicians to accept the recommendations, and adopt some action in order of minimizing the interaction risks.

Keywords:

drug interactions, medical order entry systems,
clinical decision support systems

Introduction

Errors in medicine, as well as in other human activities, occur frequently. Although most of them do not have harmful consequences, some cause injuries of varying degrees, and may even cause death. It has been reported that a fourth of those errors are related to medications [1]. Errors in medication lead to the so called preventable adverse drug events [2]. Considering the different steps in the medication cycle (prescribing, transcription, dispensing, administrating, and monitoring), it is during prescribing that almost half of the errors occur [3]. The most common mistake during this step is related to a lack of knowledge of the drug and patient information [4]. Some of the errors during the prescription step are labeled **drug-drug interactions** (DDI) [5]. A DDI occurs when one drug affects the metabolism of a second drug, thereby producing adverse effects [6]. DDI occur frequently, but most of them do not lead to adverse events. It has been estimated that only 10 to 15% of these interactions have clinical significance [7]. The occurrence of these interactions varies according to the clinical setting (inpatient, emergency, outpatient) [8]. Evidence shows that physicians do not recognize these interactions 50% of the time, and one-third of the time in serious interactions [9, 10]. Published studies and reviews indicate that **Computerized Physician Order Entry** (CPOE) that provide contextual help at the point of prescription, help in the prevention of prescribing errors and drug adverse events [11-13]. When CPOE are implemented, the clinical workflow can be affected, and generate diverse responses among the physicians using them. When asked about the potential help of CPOE in the prescription process, more than half of the practitioners agreed that they are useful [9, 14], however some studies report that doctors ignore such alerts in 57 to 95% of the time [15, 16]. One of the most important reasons for this high rate of alert overriding is evident in the literature, and is, without a doubt, due to the high rate of false positives (clinically inadequate alerts) which give rise to "alerts fatigue." This high number of incorrect alerts predisposes physicians to underestimate the CPOE as useful tools in their practice. Among the causes of the high rate of manual override of the alerts we found [8, 17]:

- Problems related with the design of the knowledge database of such systems that generate a high rate of false positives.
- Issues related to the utility of the alerts interface.
- The lack of permanent inspection of the interaction between the system and the users, in order to create cycles of continuous improvement.

The Hospital Italiano de Buenos Aires has implemented a CPOE, in the context of an electronic medical record [18], that includes an **alert system for drug-drug interactions**. This work is motivated by the fact that we have not evaluated the rate of overriding alerts by our physicians and on the problem described in the literature about these clinical decision support systems. The objective of our present work is to describe the analysis of the knowledge database of our drug-drug interaction alert system. This analysis includes annotation and purging of the knowledge database according to clinical significance, and proposes changes in its classification of recommendations and alerts visualization.

Selected for best paper award.

Table 1 - Description of fields contained in individual monographs of each DDI of the knowledge database

Drug-Drug Interaction Monographs	
Title	Both drugs (or family of drugs) participating in the interaction
Summary	Referred to the global effect of the interaction
Recommendations	Useful facts for the acting physician to resolve the case when interactions are present. Could recommend to avoid the joint prescription of the related elements, utilize an alternative drug, require complementary testing for monitoring the therapy if adopted, advise, and alert the patient, etc.
Related drugs	Extends the range of the monograph and includes pharmacological, pharmacokinetics, or chemical agents related to the drugs in question
Routes of administration	For each interaction, only the routes of administration for each particular drug are included in order to avoid the appearance of interactions alerts when the drugs are administered thru routes that do not generate an interaction (false positive)
Mechanism of interaction	Details the proposed or postulated mechanism of interaction
Significance	Code assigned to each interaction. It is based on three parameters: potential damage to the patient, frequency and predictability of occurrence, and quality of the documentation that sustains the interaction. They are classified into four levels: **Level 1:** High significance, interaction with great potential to cause damage to the patient, predictable, and frequent, and that it is well documented **Level 2:** Moderate significance, potentially damaging interaction, less predictable, less frequent, or with incomplete documentation **Level 3**: Minimal significance, interaction with low potential for damage to the patient, of variable predictability, of infrequent appearance, or that is based on little documentation **Level 4:** Without clinical significance, even thou this type of interaction can occur, the documentation is based on theoretical considerations or is not clinically significant. Also adverse effects can not be predicted
References	Includes the bibliographic references of published works sustaining the presented information

Methods

The knowledge database of the alert system was created using different sources, including clinical pharmacology textbooks, monographs on products, consultation with specialist in our institution, specific literature searching and a publication on pharmacological interactions named "Evaluation of Drug Interactions (EDI)" [19] maintained by First DataBank Inc. [20]. To build the knowledge database, the information contained in Table 1 was applied.

The EDI is organized in' 18 chapters, according to drug groups (antihypertensive drugs, narcotics, etc.) and according to its index arround 43,000 potential DDI exist. These DDI are generated from the related drugs included in each of the monographs in the EDI [20]. Our knowledge database was created mainly using this index as a guideline. Due to the smaller number of individual drug monographs of a drug-drug or drug-family interaction in comparison with the potential DDI contained in the index (1,201 vs. 43,000), the first step (done by a clinical pharmacologist), was annotating and purging each potential pair according to other sources. In addition, rounds with experienced professionals with the drug in question were also conducted. The objective was that only those interactions with a clear bibliographical and clinical background would remain in the knowledge database. Once the purging was completed, each DDI was analyzed, and a new classification of alerts was defined according to the recommendations for an action contained in the monograph. Finally all interactions were categorized either as active or passive, depending on whether they would manifest themselves actively, interrupting the prescriptive workflow or not.

Results

The first step was the **purging** of the unsubstantiated interactions based on the EDI index. Each of the 1,201 monographs in the EDI were individually analyzed, particularly the section named "Related drugs", where the 43,000 DDI included in the EDI index originated from. As a result of this analysis by a clinical pharmacologist, consultations with other sources, and revision rounds with specialists with daily experience in the use of the drugs, 39,191 DDI were discarded from the knowledge database (originally created from the EDI index as a guideline), and only 2,608 DDI were kept, each related to an individual DDI monograph. Therefore, our knowledge database was formed now with 3,809 DDI (Figure 1).

After this purging, the entries were reclassified according to their clinical significance:

- **Level 1:** 600 (High significance, interaction with great potential to cause damage to the patient, predictable, and frequent, and that it is well documented)
- **Level 2:** 1494 (Moderate significance, potentially damaging interaction, less predictable, less frequent, or with incomplete documentation)

- **Level 3:** 1653 (Minimal significance, interaction with low potential for damage to the patient, of variable predictability, of infrequent appearance, or that is based on little documentation)
- **Level 4:** 62 (Without clinical significance, even thou this type of interaction can occur, the documentation is based on theoretical considerations or is not clinically significant. Also adverse effects can not be predicted)

In addition, during the revision process, each DDI was analyzed in detail for its recommendations given to the prescribing physician. Based on this analysis, a five-domain taxonomy was created. All the recommendations contained in the alert system were then grouped according to this taxonomy (Figure 2).

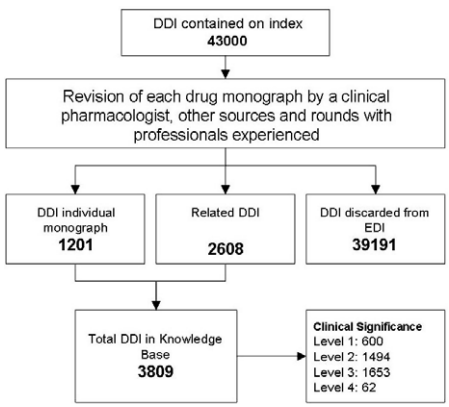


Figure 1 - Purging of the knowledge database

Recommendations taxonomy

Avoid joint use

This domain refers to those interactions that do not provide other options to the physician but to avoid both drugs in combination. This domain is the most important of all because it does not provide alternatives to the prescription. The physician must therefore justify his/her action if he/she decides to proceed and prescribe both drugs simultaneously.

Monitoring Required

This domain includes two sub domains: clinical monitoring and monitoring by complementary tests.

Clinical Monitoring includes all recommendations related to signs or symptoms that physician should inquire about during consultation and follow up with the patient to verify the effectiveness of the drug, control adverse events, or check for drug toxicity.

Some examples include:

- Clinical monitoring of desired effects:
  - Neuromuscular blockage
  - Heart frequency
  - Blood pressure
- Clinical monitoring of adverse effects:
  - Hyperglycemia
  - Gastric ulcers
  - Skin rashes
- Clinical monitoring to detect toxicity:
  - Neurotoxicity
  - Mielotoxicity
  - Cardiotoxicity

Clinical Monitoring with Complementary Tests includes laboratory tests and other studies in this category:

- Monitoring with laboratory tests:
  - Liver function tests
  - Complete Blood Count (CBC)
  - Drugs blood level
- Monitoring with other studies:
  - Electrocardiogram (EKG)
  - Electromyogram (EMG)
  - Central venous pressure

Evaluate alternative drugs

This domain includes a recommendation to search for other drugs as possible substitutions for one of the interacting pair; it also includes a support system that would provide substitution alternatives, assuring identical or similar therapeutic efficacy as the drug being replaced, and also make sure that an alternative drug will not interact with the second drug in the pair. Some examples of alternative options are:

- Acetyl salicylic acid: Acetaminophen
- Cimetidine: Famotidine/Nizatidine
- Guanetidine: Methildopa
- Erythromycin: Azytromicine

Modify administration

This recommendation does not avoid the joint administration of both drugs; however, it suggests the mode of administration, for instance:

- Space the administration of both drugs as far apart as possible in time
- Modify the dosage of one or both drugs
- Select alternative routes for administering one or both drugs
- Choose alternative pharmaceutical formats

Inform the patient

Faced with the decision of administrate an interacting pair of drugs, physicians can inform the patient about signs of alarm and other additional recommendations in order to minimize possible consequences of the interaction. Among others, they are:

- Signs of alarm related to hepatotoxicity
- Potential decrease in the contraceptive effect (evaluate alternative contraceptive methods)

- Modifications in the diet
- Signs of alarm of myolisis

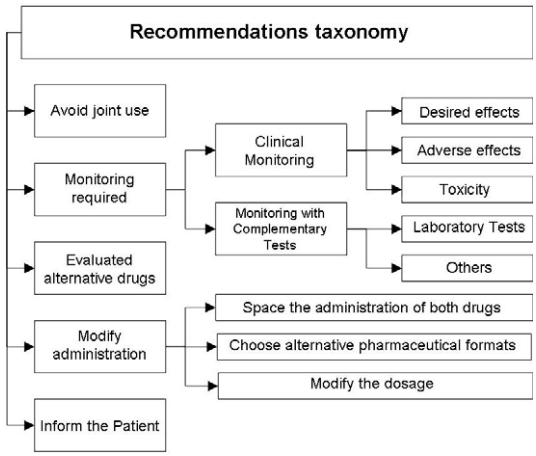


Figure 2 - Recommendations taxonomy contained in the alert system knowledge database

Alerts characteristics

Alerts are a type of clinical decision support systems and they must be intrusive, stop the workflow, and get the user's attention. Due to the fact that physicians prefer this type of alerts to be selectively used and shown in the right clinical context [21], we decided to re-categorized all DDI in the knowledge database as **active** or **passive** DDI alerts:

- **Active** DDI alerts**:** these alerts would present themselves to the physician interrupting the prescribing process, and forcing physician to explain his/her action if he/she decides to go ahead with prescription of the interacting drugs. Only 695 interactions (Table 2) are included in this category, coming from two types of DDI:
  - DDI for which the monograph recommends the avoidance of simultaneous administration. So, these DDI fall in the first domain described above ("Avoid joint use").
  - DDI that have been classified as belonging to Level 1 or High Significance are DDI with great potential to produce harm to the patient. They are frequent, and they are well documented.
- **Passive** DDI alerts: these include the remaining 3,114 interactions which allow physicians to accept the recommendations, and adopt some action in order of minimizing the interaction risks. These DDI are in the remaining domains in the DDI categories. These alerts would appear in the background, and would not interrupt the prescribing process (Table 2).

Significance	Avoid joint use Domain	Other domains
Level 1	39	561
Level 2	80	1414
Level 3	14	1639
Level 4	1	61
	Actives = 695	**Passives** = 3114

Table 2 - Re-categorization of the DDI in actives (Significance level 1 + avoid joint use domain) or passives (Significance level 2, 3 y 4 + other domains)

Discussion

The knowledge database of the support systems for decision making in the field of pharmacological prescription is generally commercially acquired, or is developed based on periodical publications. Such knowledge databases are frequently quite inclusive, putting more emphasis on the domain coverage than in the clinical relevance or the severity of adverse effects that interactions may provoke [22]. Due to this limitation, and based on the information supplied by different studies [23, 24] in which problems with the EDI are presented, we decided to undergo an analysis and subsequent purging of the knowledge database of our alert system (the EDI being an important source as guideline of our alert system as indicated above).

Clear recommendations are available as to what characteristics pharmacological alerts must meet [17, 21, 25], Therefore, the objective of the redesign of the knowledge database of our alert system was undertaken to improve the acceptance of these alerts by physicians, and minimize interruptions in the prescribing process, only leaving the most serious DDI in this group. Already there are reports that confirm that the redesign of this knowledge bases (only leaving a reduced and very specific set of alerts) has increased acceptance by physicians in general [26] and reduces the number of manual overrides of the alerts [27]. We also believe that the creation of taxonomy of recommendations related to DDI allows the physician to accept such recommendations more readily, without having his/her actions being considered as ignored alerts.

Before we implement the changes in the knowledge base of our alert system, we will conduct a study to evaluate the usability of new alerts with a group of physicians from our institution. Based on the results of this evaluation, future implementation will be decided.

References

[1] Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, et al. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. N Engl J Med. 1991 Feb 7;324(6):377-84.

[2] Nebeker JR, Barach P, Samore MH. Clarifying adverse drug events: a clinician's guide to terminology, documentation, and reporting. Ann Intern Med. 2004 May 18; 140(10):795-801.

[3] Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, Servi D, et al. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group. Jama. 1995 Jul 5; 274(1):29-34.
[4] Leape LL, Bates DW, Cullen DJ, Cooper J, Demonaco HJ, Gallivan T, et al. Systems analysis of adverse drug events. ADE Prevention Study Group. Jama. 1995 Jul 5; 274(1):35-43.
[5] Allard J, Carthey J, Cope J, Pitt M, Woodward S. Medication errors: causes, prevention and reduction. Br J Haematol. 2002 Feb;116(2):255-65.
[6] Goodman LS, Hardman JG, Limbird LE, Gilman AG. Goodman & Gilman's the pharmacological basis of therapeutics. 10th ed. New York: McGraw-Hill 2001.
[7] Seymour RM, Routledge PA. Important drug-drug interactions in the elderly. Drugs Aging. 1998 Jun;12(6):485-94.
[8] Luna D, Gonzalez Bernaldo de Quiros F, Martinez M, Gambarte ML, Morchón A, Costa G, et al. ¿Porque se ignoran los alertas sobre interacciones farmacológicas? 3er Virtual Congress of Medical Informatics - Informedica; 2004.
[9] Glassman PA, Simon B, Belperio P, Lanto A. Improving recognition of drug interactions: benefits and barriers to using automated drug alerts. Med Care. 2002 Dec;40(12):1161-71.
[10] Luna D, Hares D, Schpilberg M, Hernandez G, Soriano E, Martinez M, et al. Validación de la base de conocimiento de un sistema notificador de interacciones farmacológicas. 5to Simposio de Informática en Salud - 31 JAIIO; 2002; Santa Fe, Argentina.
[11] Bates DW, Teich JM, Lee J, Seger D, Kuperman GJ, Ma'Luf N, et al. The impact of computerized physician order entry on medication error prevention. J Am Med Inform Assoc. 1999 Jul-Aug;6(4):313-21.
[12] Poikonen J, JM L. Medication-Management Issues at the Point of Care. J Healthc Inf Manag. 1999;13(2):43-52.
[13] Teich JM, Merchia PR, Schmiz JL, Kuperman GJ, Spurr CD, Bates DW. Effects of computerized physician order entry on prescribing practices. Arch Intern Med. 2000;160(18):2741-7.
[14] Magnus D, Rodgers S, Avery AJ. GPs' views on computerized drug interaction alerts: questionnaire survey. J Clin Pharm Ther. 2002 Oct;27(5):377-82.
[15] Payne TH, Nichol WP, Hoey P, Savarino J. Characteristics and override rates of order checks in a practitioner order entry system. Proc AMIA Symp. 2002:602-6.
[16] Weingart SN, Toth M, Sands DZ, Aronson MD, Davis RB, Phillips RS. Physicians' decisions to override computerized drug alerts in primary care. Arch Intern Med. 2003 Nov 24;163(21):2625-31.
[17] van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc. 2006 Mar-Apr;13(2):138-47.
[18] Luna D, Otero P, Gomez A, Martinez M, García Martí S, Schpilberg M, et al. Implementación de una Historia Clínica Electrónica Ambulatoria: "Proyecto ITALICA". 6to Simposio de Informática en Salud - 32 JAIIO; Buenos Aires, Argentina; 2003.
[19] Zucchero F, Hogan M, Sommer C, eds. Evaluations of drug interactions. St. Louis, Missouri, USA: First DataBank Inc. 2005.
[20] First DataBank Inc. Evaluations of Drug Interactions (EDI). 2005 [accessed 10 Oct 2006]; Available from:http://www.firstdatabank.com/products/evaluations/
[21] Krall MA, Sittig DF. Clinician's assessments of outpatient electronic medical record alert and reminder usability and usefulness requirements. Proc AMIA Symp. 2002:400-4.
[22] Reichley RM, Seaton TL, Resetar E, Micek ST, Scott KL, Fraser VJ, et al. Implementing a commercial rule base as a medication order safety net. J Am Med Inform Assoc. 2005 Jul-Aug;12(4):383-9.
[23] Abarca J, Malone DC, Armstrong EP, Grizzle AJ, Hansten PD, Van Bergen RC, et al. Concordance of severity ratings provided in four drug interaction compendia. J Am Pharm Assoc (Wash DC). 2004 Mar-Apr;44(2):136-41.
[24] Malone DC, Abarca J, Hansten PD, Grizzle AJ, Armstrong EP, Van Bergen RC, et al. Identification of serious drug-drug interactions: results of the partnership to prevent drug-drug interactions. J Am Pharm Assoc (Wash DC). 2004 Mar-Apr;44(2):142-51.
[25] Feldstein A, Simon SR, Schneider J, Krall M, Laferriere D, Smith DH, et al. How to design computerized alerts to safe prescribing practices. Jt Comm J Qual Saf. 2004 Nov;30(11):602-13.
[26] Spina JR, Glassman PA, Belperio P, Cader R, Asch S. Clinical relevance of automated drug alerts from the perspective of medical providers. Am J Med Qual. 2005 Jan-Feb;20(1):7-14.
[27] Shah NR, Seger AC, Seger DL, Fiskio JM, Kuperman GJ, Blumenfeld B, et al. Improving acceptance of computerized prescribing alerts in ambulatory care. J Am Med Inform Assoc. 2006 Jan-Feb;13(1):5-11.

Address for correspondence

Dr. Daniel Luna: daniel.luna@hospitalitaliano.org.ar Área de Informática Médica. Hospital Italiano de Buenos Aires. Gascón 450. Ciudad Autónoma de Buenos Aires. Argentina. (C1181ACH). Tel/Fax:+54-11-4959-0507

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Closing the Loop : Bringing Decision Support Clinical Data at the Clinician Desktop

Claudine Bréant[a], Francois Borst[a], René Nkoulou[b], Olivier Irion[c], Antoine Geissbuhler[a]

[a]*Service of Medical Informatics, University Hospitals of Geneva, Geneva, Switzerland*
[b]*Division of Cardiology, University Hospitals of Geneva, Geneva, Switzerland*
[c]*Division of Obstetrics, University Hospitals of Geneva, Geneva, Switzerland*

Abstract

We describe the development of an inquiry office to bridge the gap between clinician needs for decision support systems and readily available large quantities of integrated clinical data. With this link, an information feedback mechanism is implemented that closes the loop of information flow by bringing decision support information from the data warehouse at the clinician desktop. As a result, and as a new DRG cost reimbursement system has been introduced, we have provided the heads of over 30 medical services with an intranet web-based application to access patient encoding of diagnoses, procedures, and DRGs of their respective service. The inquiry office has also developed a query service to process specific requests. It has implemented the automatic screening of patient clinical data of past and current hospitalizations in order to select cases for multiple studies, research, and teaching projects. The purpose of this clinical data warehouse and its information feedback process is to offer a coherent, comprehensive, and reliable return of information to improve decision making, to enable research projects, and to facilitate statistical outputs.

Keywords:

data warehouse, hospital information system, feedback, decision support system

Introduction

With vast amounts of high quality clinical data stored in various databases, the challenge today resides in extracting dynamic information from static data stores. Ultimately, the goal is to provide intelligent and relevant guidance at the point of decision making towards the delivery of better clinical health services. This translates into the integration of information feedback into clinical practice, medical management, research and teaching projects.

Although, it is frequently discussed that electronic patient record (EPR) systems have a tremendous potential to improve care and performance measurements, few developments have been made in the attempt of bringing practice feedback at the clinician desktop [1].

Until now, the prime focus has been to design data warehouses to tackle heterogeneity of medical databases and lack of connection between clinical systems [2][3][4]. The data warehouse technology has become widely spread in the clinical area and health institutions report on the value of integrating their data into a single repository [5][6]. A data warehouse can be simply defined as a copy of transactional data specifically structured, integrated, and organized for complex querying, data analysis, and decision support applications [7].

However, clinicians, medical managers, and researchers understand well their problematic but do not necessarily master the system query languages and the data structures. Today, the full exploitation of clinical data warehouses is impeded by the lack of user-friendly interfaces integrated at the clinician desktop. Furthermore, truly effective methods of feedback from a clinical data warehouse have yet to be found [1]. Needs in terms of access to a clinical data warehouse have been evaluated in the setting of an academic healthcare center and are three folds:

1. automatic patient screening with notification of cases matching a pre-defined criteria (i.e. similar cases retrieval). This mechanism is necessary to both accurately filter patient cases and avoid the tedious task of manually reviewing patient records for a particular study, research, or teaching project [8].
2. on-line dynamic navigation to explore views displaying both indicators (i.e. aggregated data) and detailed patient data. In a multi-dimensional data warehouse, indicators can be calculated across patient records and medical specialties along several axes of analysis [7]. To build a strong analysis, the end-users need to access these indicators as well as the underlying detailed data.
3. data mining and statistics functionalities for automatically searching large volumes of clinical data to extract new patterns representing useful information [9][10][11].

An additional important requirement is the patient data confidentiality that must be enforced according to the institution regulations.

In this study, we describe the development of an efficient inquiry office to bridge the gap between clinicians' needs

and readily available large quantities of integrated clinical data. The paper is organized as follows; materials and methods are described first with a brief review of the overall data workflow architecture, including more specifically the data warehouse content and data model. The focus is then on the information feedback process that was lately added to bring decision support information from the data warehouse to the clinical desktop. With this new link, an information feedback mechanism is implemented that closes the loop of information flow as depicted in Figure 1. Finally, results are presented for : 1) an intranet web-based application for medical activity analysis linked to a new cost reimbursement system, and 2) for a query service for a customized access to the clinical data warehouse.

Materials and methods

The data workflow architecture

The electronic patient record (EPR) organizes the acquisition and the visualization of patient clinical data during the care production workflow. Furthermore, this data is managed by the various independent clinical care systems such as for instance the radiology information system (RIS), the admission/discharge/transfer (ADT) system, and the laboratory system. In this setting, the University Hospitals of Geneva (HUG) have been accumulating large amounts of clinical data for various domains of clinical activity, embedding knowledge and experience produced during the patient care processes. However, to improve and evaluate clinical practice, information processing must go beyond basic automation (i.e. patient oriented access only) and convert EPR data into aggregated, multidimensional information. Hence, following the tendency in other industries, it requires the integration of patient clinical data into a single repository.

Table 1 – Content of the Archimed data warehouse.It currently includes seven main data sources from the HIS.

Archimed Clinical Data Marts	**HIS components (*data source names*) and years covered**	**Facts (millions)**
ADT	1990 (*Impact*) ->2005 (*DPA*)	10M
Diagnoses (ICD codes)	1990 (*PDP*)	2.8M
Procedures (CHOP codes)	1990 (*PDP*)	1M
DRG codes	2005 (*APDRG*)	0.5M
Laboratory	1993 (*Unilab*)->2005 (*Unilab2*)	75M
Clinical data questionnaires	2002 (*DPI-Form*)	10M
Radiology exams	1990 (*Unimage*)->2004 (*Xplore*)	2M

The Archimed clinical data warehouse database has succeeded in procuring an integrated and coherent view of a wide variety of medical and administrative data, all centered on patient care processes, acquired from multiple

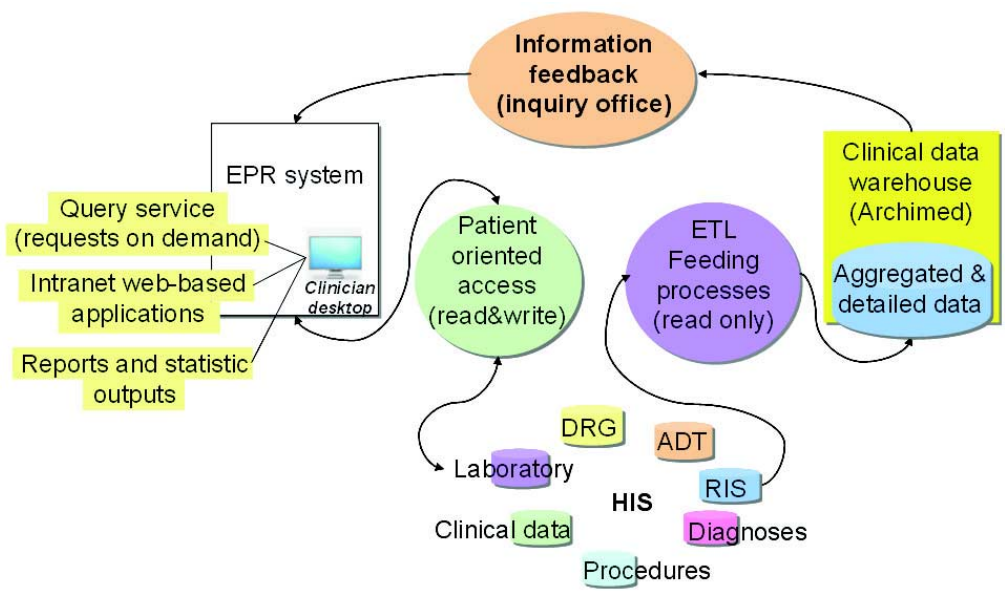


Figure 1 – The data workflow architecture. The EPR system provides a read/write access to clinical data that is patient oriented only. ETL processes are run daily to extract, transform, and integrate HIS data into a single multi-dimensional database (Archimed). The information feedback process allows the return of decision support information towards the clinician desktop.

independent and heterogeneous systems part of the HIS [12]. Extract/Transform/Load (ETL) processes are run daily to refresh data. It currently stores patient clinical data since the early 90's for a total of over 750'000 patients, 1.7M hospitalization stays, and 2M out-patient visits. More specifically, the Archimed database includes around 100 millions patient facts spread over 7 data marts, as listed in Table 1. As an example, the ADT data mart stores detailed patient stay administrative information consisting in 10 millions of patient facts. It was fed by the *Impact* ADT system until the end of 2004, and then by the upgraded *DPA* system from 2005.

Archimed is designed according to the data warehouse bus (DWB) architecture and the multi-dimensional model advocated by Kimball [7]. In this architecture, the data warehouse is composed of a set of data marts connected to each other through some specific dimensions called conformed dimension tables. The conformed dimension tables store key data on which each care process relies and on which the whole institution agrees and complies. In the case of Archimed, the conformed dimension tables have been identified at an early stage of the design and model the axes of analysis of the data warehouse. These are the patient administrative data, the episode of care description, the medical services, the care units of the hospital, as well as the care centers. This architectural design builds the foundation for a simple and powerful querying of the patient medical data as if it were initially part of the same database [12]. Figure 2 illustrates the Archimed data model. Black dots show how the fact tables are linked to the conformed dimensions tables.

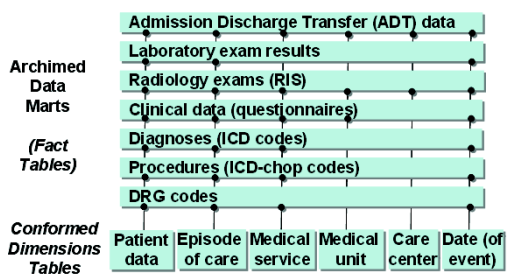


Figure 2 – The Archimed clinical data warehouse bus architecture

Based on the DWB architecture, each data mart or domain of activity has been progressively integrated into the Archimed system, breaking the overall implementation task to manageable proportions. This architecture also provides scalability to the system, facilitating the future addition of new data marts.

The information feedback process

A new inquiry office has been created to go beyond the production of simple static reports and to fully leverage the value of the Archimed data warehouse. The inquiry office is responsible for implementing the information feedback process to bring decision support data at the clinician desktop. It fulfills the following tasks:

- Act as a mediator between the data warehouse system and the clinicians needs,
- Develop both computerized tools and their integration into the clinician workflow,
- Enforce procedures regarding patient data confidentiality, according to our institution regulations.

The restitution of information is currently performed according to the two following methods: 1) intranet web-based applications, 2) a query service for a customized access to the data warehouse.

On-line access through intranet web-based applications

We have utilized the intranet web-based technology to facilitate the navigation into a pre-defined subset of data at different levels of granularity. Views of clinical information, possibly combining the data from several data marts are prepared from the data warehouse. These are available on-line at the desktop of the end-user through intranet web-based applications. The user can define and parameterize views by, for instance, choosing a period of interest or by specifying a patient population. Different levels of granularity can be presented to convey the information. These range from aggregates to summarize data at various levels of interest, to detailed patient clinical data. *Drill-down* and *roll-up* links allow to navigate between views to explore more detailed data, or on the contrary to analyze more synthetic information.

Furthermore, the intranet web-based applications managed by the inquiry office are connected to the hospital access right management system hence providing the level of data confidentiality required by the institution.

A query service for customized access to the data warehouse

We have defined a new environment to automatically screen patient data according to a user-defined criterion. As described above, our data model gives us the ability to easily build complex criterions involving data from as many clinical activities as needed. Such queries take only a few seconds to a few minutes to run against the whole database. The basic result obtained is a list of patient cases. According to users needs, this list can be enriched in a second step with an extraction of partial or total clinical data for the patients of the initial list. Finally, results are communicated either simply by e-mail, or through intranet web-based views made available within the EPR environment.

The processing of a new request is as follows. The clinician first contacts the inquiry office to communicate a new request. Proper authorization to access the data is checked according to the HUG patient confidentiality regulations. Then, the inquiry office builds the queries that access the data warehouse, and results are carefully validated with the clinician. Some parameters are set such as the frequency and the duration of the hospital patient population screening.

Results

Statmed : a desktop tool to understand the implications the new DRG reimbursement system at the HUG

As in many countries, the reimbursement system in Switzerland for acute inpatient care has recently transferred from a standard daily reimbursement to a complete diagnosis-related group (DRG) reimbursement system. The DRG system classifies hospital cases into one of approximately 900 groups, expected to have similar hospital resource use. It is now part of the HUG payment system since 2006. DRGs are assigned by a grouper program mainly based on ICD diagnoses, procedures, age, sex, and the presence of comorbidities. Reimbursement is calculated from a relative weight of a procedure multiplied by a base rate.

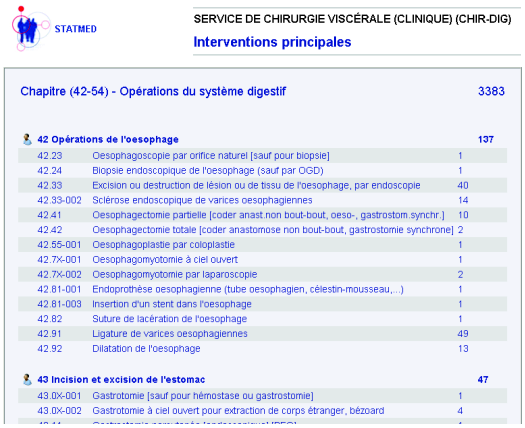


Figure 3 - View of the procedures performed by the division of visceral surgery and corresponding number of cases. The patient icon allows to drill down to more detailed patient medical information.

This is a major change for the HUG and some of the implications are unknown. In that case, information feedback is of paramount to accompany the change and help users to get insights and understanding of the new reimbursement system. Consequently, we have provided the heads of over 30 medical services with an intranet web-based application available at their desktop to access both patient detailed and summarized data related to the encoding of diagnoses, procedures, and DRGs of their respective service.

The user first selects a period of interest and a medical service (i.e. division of visceral surgery). Then, he can choose to view the number of diagnoses, procedures (Figure 3), or DRGs (Figure 4) coded for patients who completed their stay within the period of interest in the selected medical service.

For further analysis, the user can select a DRG code and drill-down to the encoding of the underlying diagnoses and procedures to study the link between the medical activity and the financial outcome. Finally, from any dashboard displaying codes of diagnoses, procedures, or DRGs, the user can retrieve the corresponding list of patient identification numbers and connect to the EPR system to have a view of the entire record file to explore other details regarding the patient hospitalization stay. For instance, when encountering an intriguing anomaly in the data, it is then necessary to access the full patient record to determine the exact cause.

In addition to the clinicians, this program is used by the administrative staff of about 15 members, responsible for the ICD encoding of the patient diagnoses and procedures. They regularly use the program to perform quality control verifications, to detect possible errors in code attribution, and to monitor their activity.

Figure 4 - View of the DRG codes for patients hospitalized in the division of visceral surgery. Icons on the left part of the screen can be used to drill down to a more detailed description of underlying patient cases, diagnostic and procedure codes.

The current release of Statmed merges data from four different data marts: Diagnostic and Procedure ICD Encoding, DRG codes, and ADT administrative data. It is a Java based program that automatically generates Structured Query Language (SQL) queries to retrieve data, to build aggregate and then to display dashboards presenting the data.

Automatic patient screening and notification of selected cases

The inquiry office is regularly called for support by clinicians for implementing the automatic screening of patient clinical data of past and current hospitalizations in order to select cases for multiple studies, research, or teaching projects. As an illustration, we have listed below some examples of patient screening requests among about eighty inquiries processed in 2006 at the HUG:

- The european multicenter Micado randomized controlled trial which enrolls women with preterm premature rupture of membranes between 28 and 32 weeks (division of obstetrics).
- The multicenter AMIS registry and the MIDAS study collect data on patients with acute coronary syndrome in Switzerland (division of cardiology).

- Study regarding guideline and management of adult asthma in Switzerland (division of general internal medicine).
- Study of children with cystinuria (division of nephrology).
- Extraction of clinical data for the calculus of the APACHE and SAPS scores for the patients of the intensive care unit (division of intensive care).
- Epidemiological study of hip fractures among the elderly population in Geneva's county (Switzerland) (division of geriatrics).
- Two-years follow-up of intensive care patients (division of intensive care).
- Study of the obstetrical outcome after treatment of cervical dysplasia (division of obstetrics).

Regarding qualitative results, and more particularly in the case of the AMIS registry, it has been established that a retrospective database can be used to analyze clinical practice and furthermore participate in the modification of the therapeutic behavior of the physicians [13].

Conclusion

In this work, we have demonstrated the feasibility of a mechanism to bring dynamic data prepared from a hospital-wide clinical data warehouse to the clinician desktop. The results obtained so far pave the way for the development of more client interfaces depending of specific needs. The web-based technology allows these applications to be quickly developed and required no deployment effort. Hence, the data warehouse is not any longer a stand-alone system but becomes fully integrated into the clinician routine workflow.

To fully leverage the data warehouse towards a usable information system, we have created an inquiry office including methodology and tools in order to assist users access data. Graphical user interfaces and web-based technology are used to provide on-line meaningful views of clinical information. As of today, clinicians are using the system for analysis of the new HUG cost reimbursement scheme and for monitoring their medical activities. This system necessitates very few training. Specific requests regarding case studies, the retrieval of similar cases, and statistics are processed on demand by the inquiry office. About eighty requests have been processed in 2006. We have defined a new environment to automatically screen patient data according to a user-defined criterion, and to bring results at the clinician desktop.

The long-term objective of this work is to have this system becoming an integrated part of the clinician workflow for quantitative measurable results. Existing tools will need to evolve, and new developments will be initiated to follow the users needs. For instance, we have tested the coupling of Archimed with data mining tools such as the Weka open source toolkit and Clementine®. A few data mining projects have been initiated in the area of detection and surveillance of nosocomial infections, and to study patient readmissions.

References

[1] Grant A, Moshyk A, Diab H, Integrating feedback from a clinical data warehouse into practice organization. Int J Med Informatics, 75:232-239, 2006.

[2] Ledbetter CS, Morgan MW, Toward best practice: leveraging the electronic patient record as a clinical data warehouse J Healthc Inf Manag. 2001 Summer;15(2):119-31.

[3] Ebidia A, Mulder C, Tripp B, Morgan MW. Getting data out of the electronic patient record: critical steps in building a data warehouse for decision support. Proc AMIA Symp. 1999;:745-9.

[4] Einnbinder JS, Scully KW, Pates RD, Schubart JR, Reynolds RE. Case study: a data warehouse for an academic medical center. J Health Inf Manage 15 (2001) 165-175.

[5] Wisniewski MF, Kieszkowski P, Zagorski BM, Trick WE, Sommers M, Weinstein RA. Development of a clinical data warehouse for hospital infection control. J Am Med Inform Assoc. 2003 Sep-Oct:10(5):454-62.

[6] McGregor C,Bryan G, Curry J, Tracy M. The E-baby data warehouse: a case study. 35th Annu Hawaii Intl. Conf. on System Sciences, 2002.

[7] Kimball R, Reeves L, Ross M, Thornthwaite W (1998) The data warehouse lifecycle toolkit.Wiley, New York.

[8] Stephen R, Boxwala A, Gertman P. Feasibility of using a large Clinical Data Warehouse to automate the selection of diagnostic cohorts. AMIA Annu Symp Proc. 2003;:1019.

[9] Cao H, Markatou M, Melton GB, Chiang MF, Hripcsak G. Mining a clinical data warehouse to discover disease-finding associations using co-occurence statistics. AMIA Annu Symp Proc. 2005;:106-10.

[10] Wan TT. Healthcare informatics research: from data to evidence-based management. J Med Syst. 2006 Feb;30(1): 3-7.

[11] Silver M, Sakata T, Su HC, Herman C, Dolins SB, O'Shea MJ. Case study: how to apply data mining techniques in a healthcare data warehouse. J Healthc Inf Manag. 2001 Summer;15(2):155-64.

[12] Breant C, Thurler G, Borst F, Geissbuhler A. Design of a Multi Dimensional Database for the Archimed Data Warehouse.Stud Health Technol Inform. 2005;116:169-74.

[13] Fassa A-A, Urban P, Radovanovic D, Duvoisin N, Gaspoz JM, Stauffer J-C, Erne P; Trends in reperfusion therapy of ST segment elevation myocardial infarction in Switzerland: six year results from a nationwide registry. Heart 2005;91: 882-888

Address for correspondence

Claudine Bréant PhD,
Service of Medical Informatics,
University Hospitals of Geneva (HUG),
1211 Geneva, Switzerland.
Email: claudine.breant@sim.hcuge.ch

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Are Clinicians' Information Needs and Decision Support Affected by Different Models of Care? Experimental Study

Vitali Sintchenko[a,b], Tim Usherwood[b], Enrico Coiera[a]

[a]*Centre for Health Informatics, University of New South Wales,*
[b]*Western Clinical School, University of Sydney, Sydney, Australia*

Abstract

This study explores task- and healthcare model-specific differences in clinicians' information needs which can affect the uptake of decision support. Results of a web experiment involving 104 general practitioners are presented. Respondents indicated that guidelines were the most important source of information with almost equal weighting for acute, chronic and preventive care. A patient's quality of life was identified as the most important determinant of decision-making in all three models of care. Risk assessment tools and information about outcomes were more valuable (P<0.05) for chronic and preventive care than for acute cases. The participants accessed electronic risk assessment tools in 54%, 45% and 81% of acute, chronic and preventive care scenarios, respectively. Participants estimated that the electronic decision support would have a significantly higher impact in preventive care than in chronic or acute care settings (P=0.01). The differences in the information needs of clinicians related to different care models have to be considered in the design of clinical decision support systems. Systems that target preventive model decisions may have higher adoption and impact.

Keywords:

information needs; information systems; clinical decision-making

Introduction

Lack of knowledge about the information needs of clinicians has been identified as one of the major reasons for the slow uptake of decision support tools in clinical practice [1,2,3]. Evidence suggests that the type of decision task affects the effectiveness of electronic decision support systems (EDSS). The provision of computerised decision support improves the quality of decision-making and outcomes of patients with acute illnesses in a hospital setting. However, the impact of EDSS on care decisions for patients with chronic conditions or genetic risks is less certain [4,5]. Despite the increasing availability of genetic testing and the emergence of new programs for shared genetic risk assessment [6,7], knowledge of clinician's predictive testing and EDSS use remains limited [8]. A large variety of attitudes and beliefs influence clinical decision-making by clinicians [5]. Our previous research points to task- and healthcare model-specific differences in their information needs, which can affect the uptake of EDSS [9]. To test this hypothesis, three modules of decision support were developed and a web-based experiment was designed to aid clinical risk assessment for acute, chronic and preventive care models. The objectives of this experiment were to explore (a) task specific information needs related to acute, chronic and preventive care, and (b) the free-willed use of computerised decision support tools by clinicians performing the above tasks.

Methods

Study population and design

We surveyed a convenience sample of Australian general practitioners to examine the differences in their information needs and beliefs about the role of EDSS in acute, chronic and preventive care. Participants were recruited among Fellows and Trainees of the Royal Australian College of General Practitioners through the Quality Assurance & Continuing Professional Development Programme. Participation was voluntary and no monetary incentives were offered. Participants were randomised to the EDSS group and the unaided decision-making group and asked to review three clinical scenarios in a random order to avoid learning effect as well as to answer case-specific questions (Figure 1). The first group was provided with access to decision support modules at the time of answering the questions. The modules provided probabilistic patient-specific information about future clinical outcomes. Clinicians were not specifically informed that their use of EDSS modules would be logged. The second group only saw the decision support modules after answering case-specific questions at the end of the survey. The questions were designed using a 5-point Likert scale to identify potential factors affecting EDSS acceptance as well as specific usage patterns on the part of clinicians. At the end of the survey each participant was provided with feedback regarding (a) the distribution of answers from the group of participants, including their own answers, and (b) the impact of the use of decision support tools on this process.

The experiment was conducted between April and August 2006 as an online exercise with interactive decision sup-

port modules accessible from a personal computer over a secure Internet connection to the University of NSW server. This survey can be accessed from: http://129.94.108.23/DSSSurvey/index.jsp

Decision support modules

Three risk assessment aids were built using published scores based on demographic, clinical and behavioural information that is routinely collected when taking a history (Copas, 2002). Community-acquired pneumonia decision support provided risk assessment based on the Pneumonia Severity Index [10] and the genital herpes risk assessment tools was based on local Australian data. The Breast Cancer Risk Assessment module was based on probabilities obtained from two recently validated models: the BOADICEA model of genetic susceptibility to breast cancer from the University of Cambridge [11] and BRCARPO model of genetic testing of BRCA1/BRCA2 and prevalence of breast cancer [12] (Figures 2 and 3). Scenarios of community-acquired pneumonia, genital herpes and breast cancer were designed specifically to represent, respectively, acute, chronic and preventive models of care because of their relative differences in the acuity and complexity of clinical decision-making, as well as the utility of shared decision-making and decision outcomes.

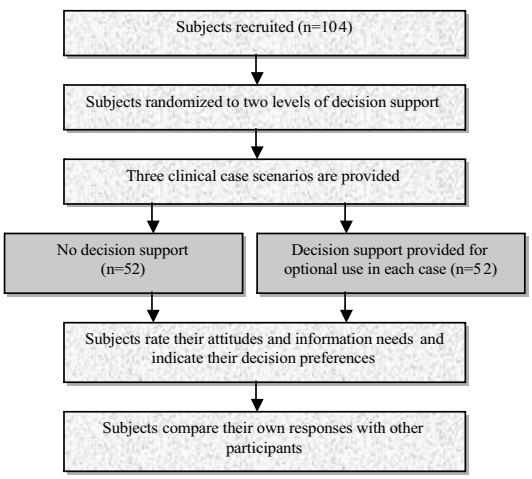


Figure 1 - An online experiment design

Outcome measures

Using a log-file, the participants' choice to use a decision support tool was monitored. The attitudes and self-reported information needs of subjects were recorded. The impact will be assessed using one-way analysis of covariance (ANOVA) and chi-square statistics. Statistical significance was set at $P<0.05$.

Results

Characteristics of participants

104 general practitioners completed the experiment. Twenty-five of respondents (25%) also treated residents of aged care facilities and 15 (14%) served as sessional hospital medical officers. All respondents were accredited practitioners representing all States and Territories of Australia. Half of the participants (n=53) indicated that they use a computer in for keeping medical records and electronic prescribing. General practitioners familiar with computers were equally distributed in the two study groups.

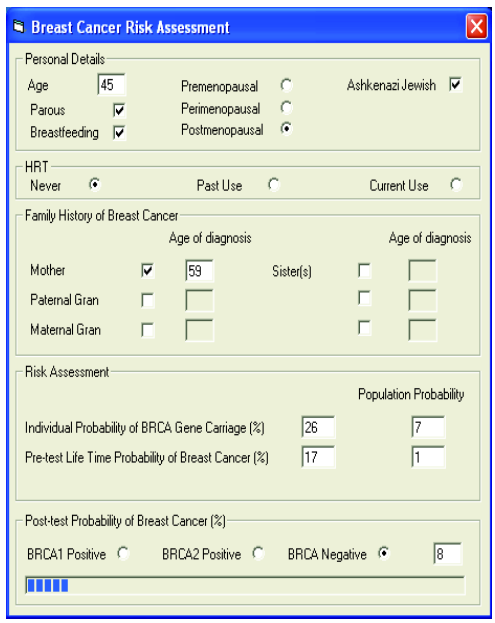


Figure 2 - Snapshot of the decision support module for the risk assessment of breast cancer

Information needs for acute, chronic and preventive care

Overall, participants were positive about the value of probabilistic EDSS in primary care. All clinicians rated clinical guidelines, condition-specific risk calculators and computerised decision support as important sources of evidence (Table 1). The potential impact on a patient's quality of life and his or her satisfaction were the two most important determinants of clinicians' decision-making in all three models of care. Respondents indicated that clinical guidelines were the most important source of information with almost equal weighting for acute, chronic and preventive care.

Condition-specific risk calculators and information about outcomes of care for previous patients with similar problems were more valuable ($P<0.05$) for chronic and preventive care situations than for acute cases (Table 1). When two arms of the study were compared, the participants in the decision support arm of the experiment accessed electronic risk assessment tools in 54%, 45% and 81% of acute, chronic and preventive care scenarios,

*Table 1 - Attributes affecting the quality of clinical decision-making for different models of care, mean scores**

Statement	Acute care	Chronic care	Preventive care	P**
The importance in decision-making				
Compliance with clinical guidelines	1.793	1.526	1.758	NS**
Cost to taxpayer	2.414	2.069	2.379	NS
Impact on quality of life	1.448	1.345	1.276	NS
Patient satisfaction	1.448	1.241	1.310	NS
Access to clinical guidelines	1.862	1.757	1.758	NS
Access to condition-specific risk calculators	2.345	2.034	1.862	0.04
Information about outcomes of care of previous patients with similar problems	2.276	1.793	1.861	0.05
Information essential for optimal decision-making				
Document patient history	1.621	1.413	1.552	NS
Clinical protocols	2.034	1.758	1.621	NS
Cost of care to a patient	2.793	1.897	2.931	0.001
Information most often lacking				
Document patient history	2.414	2.414	2.931	0.03
Clinical protocols	2.655	3.310	1.828	0.002
Cost of care to a patient	2.517	2.655	1.931	0.02

* Items rated on scale of 1 (Strongly agree) to 5 (Strongly disagree). Statistics were calculated by ANOVA.
** NS, not significant.

respectively. The majority of general practitioners who used decision support tools also felt that EDSS could have a high impact on the quality of decision-making in different models of care (Figure 3).

Information gaps in acute, chronic and preventive care

When specifically asked about information types essential for optimal decision-making, participants agreed that a documented patient history, relevant test results and clinical protocols were the most relevant pieces of information.

Respondents felt that the cost of care to a patient was also a contributor to chronic care decisions and was more important for chronic care than for other models P=0.001). When asked to identify the specific types of information most often lacking at the time of clinical decision-making, respondents indicated that a relevant patient history and test results were more often lacking in acute and chronic care. In contrast, clinical protocols (P=0.002) and cost information (P=0.02) were more often lacking in preventive care (Table 1).

Participants estimated that the electronic decision support would have a significantly higher impact in preventive care than in chronic or acute care settings (P=0.01) (Figure 4).

Discussion

This study showed that perceived information needs between acute, chronic and preventive care models in general practice were similar. The main differences in information needs in these settings were related either to the availability of information or to access to decision support.

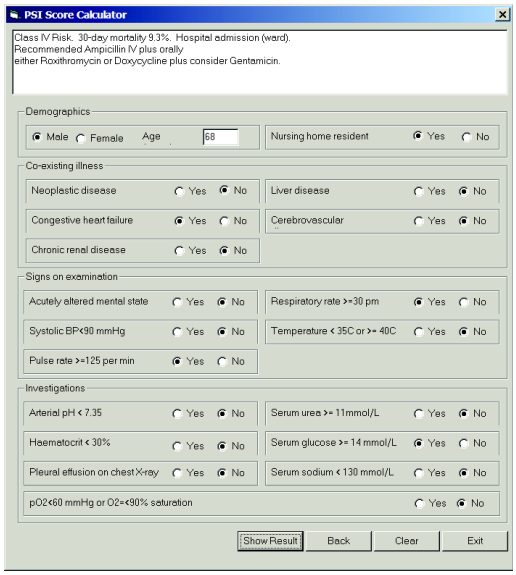


Figure 3 - Snapshot of the decision support module for the risk assessment of community-acquired pneumonia

While the lack of documented patient history and test results has been judged as more important in acute and chronic care, finding appropriate clinical protocols and information about the cost of care was of higher priority in preventive care. Furthermore, medical professionals indicated that the successful information searching has a more significant impact on clinical decision-making in preventive care than in other setting. It could reflect differences in the goals of management or variations in risk preferences for different scenarios shown to influence clinical decisions [5]. Although physician uncertainty has been cited as a cause of practice pattern variations [5,9], differences in physicians' objectives concerning the goals of treatment and management in different models of care have received less attention [13]. However, these differences imply that clinicians may use different information seeking strategies for acute, chronic and preventive care tasks.

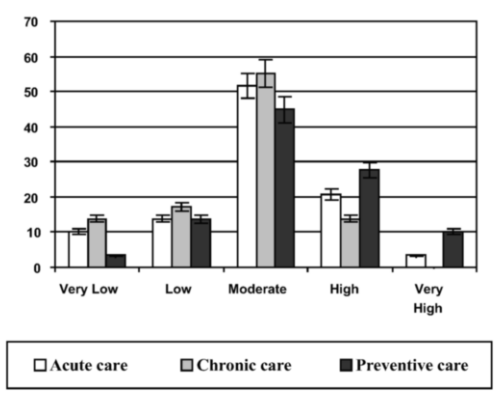


Figure 4 - Distribution of the perceived impact of information provided by decision support tools on the quality of clinical decision making in different models of care. Estimates: very low=0-20%; low=21%-40%; moderate=41%-60%; high=61%-80%; very high=81%-100%.

Many decision support systems are based on electronic guidelines or protocols. However, this experiment gathered further evidence to support the suggestion that the impact of computerised decision support exceeds that of guidelines [9]. Our results also suggest that the uptake of decision support may be affected by differences in acute, chronic and preventive care tasks. Our participants more frequently used a decision support tool for the breast cancer risk assessment than for acute or chronic care conditions. Our findings about the use of condition-specific risk calculators using pathology testing data are of particular significance as diagnostic tests account for about 25% of ambulatory health care costs, with 80% of all health expenditures directed by physicians [14].

Previously undocumented facts have emerged from our experiment: a higher utility and uptake of decision support for preventive decision tasks than for acute care decisions; a gap in decision support for preventive medicine; significant differences in the perceived impact of information on the clinical decision-making process for acute, chronic and preventive models of care.

These conclusions should be interpreted in light of the limitations in the study design. First, the experiment relied on self-reported behaviour without verification that clinicians actually practice in the manner described. Our study was experimental and did not take into account operational issues concerning the application of EDSS in a clinical setting. It is also possible that clinicians in the intervention arm used EDSS relatively often because they felt they were participating in a new study. Second, we used a convenience sample of primary care practitioners. Participation was voluntary and clearly the study may be biased towards those who felt more comfortable with electronic decision support. However, clinical vignettes as a method for measuring the competence of physicians and the quality of their actual practice have been validated [15] and interactivity in the web-based questionnaires increased compliance [16,17]. Lastly, we surveyed general practitioners. Specialist clinicians are likely to differ from primary care providers in decision-making styles and information needs [18]. However, this has the advantage of representing the point of view of professionals who see the whole spectrum of problems at the front end of health care delivery and do not necessarily represent "early adopters" of new concepts.

In conclusion, the differences in the information needs of clinicians related to different care models have to be considered in the design and implementation of clinical decision support systems. Systems that target preventive model decisions may have higher adoption and impact. These findings discover the relative value of different types of information needed for the optimisation of clinical decision-making in primary care and identify key strate-

gies for the design and implementation of successful EDSS for genomic medicine.

Acknowledgments

VS was supported by the Australian National Institute of Clinical Studies. Authors are indebted to Ken Nguyen for his help in the design of the web questionnaire and also thank the Quality Assurance & Continuing Professional Development Programme of the Royal Australian College of General Practitioners and the GPs who participated in the survey.

References

[1] Sim I, Gorman P, Greenes RA, Haynes RB et al. Clinical decision support systems for the practice of evidence-based medicine. J Am Med Inform Assoc 2001;8:527-34.

[2] Pluye P, Grad RM, Dunikowski LG, Stephenson R. Impact of clinical information-retrieval technology on physicians: A literature review of quantitative, qualitative and mixed methods studies. Intern J Med Inform 2005;74(9):745-68.

[3] Dreiseitl S, Binder M. Do physicians value decision support? A look at the effect of decision support systems on physician opinion. Artif Intell Med 2005;33:25-30.

[4] Bates D, Kuperman G, Wang S, Ganghi T, Kittler A, Volk L, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc 2003;10:523-30.

[5] D'Alessandro DM, Kreiter CD, Peterson MW. An evaluation of information-seeking behaviours of general practitioners. Pediatr 2004;113:64-9.

[6] Ravdin PM, Siminoff LA, Davis GJ, Mercer MB, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol 2001;19:980-91.

[7] Baum M, Ravdin PM. Decision-making in early breast cancer: guidelines and decision tools. Eur J Cancer 2002;38:745-9.

[8] Knottnerus JA. Community genetics and community medicine. Fam Pract 2003;20:601-6.

[9] Sintchenko V, Coiera E. Decision complexity affects the extent and type of decision support use. In: Bates D (Ed). Proc Annual AMIA Symp, Washington DC, 11-15 Nov 2006; pp.724-8.

[10] Dean NC, Suchyta MR, Bateman KA, Aronsky D, Hadlock CJ. Implementation of admission decision support for community-acquired pneumonia: a pilot study. Chest 2000;117:1368-77.

[11] Antoniou AC, Pharoah PPD, Smith P, Easton DF. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. Br J Cancer 2004;91:1580-90.

[12] Berry DA, Iversen ES, Gudbjartsson DF, Hiller EH, et al. BRCARPO validation, sensitivity of genetic testing of BRCA1/BRCA2, and prevalence of other breast cancer susceptibility genes. J Clin Oncol 2002;20:2701-12.

[13] Flach SD, Longenecker JG, Tape TG, Bryan TJ, Parenti C, Wigton RS. The relationship between treatment objectives and practice patterns in the management of urinary tract infections. Med Dec Making 2003;23;131-9.

[14] Birbeck GL, Gifford DR, Song J, Belin TR, Mittman BS, Vickrey BG. Do malpractice concerns, payment mechanisms, and attitudes influence test-ordering decisions? Neurol 2004;62;119-21.

[15] Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients and chart abstractions. J Am Med Assoc 2000;283:1715-22.

[16] Emery J, Walton R, Murphy M, Austoker J, et al. Computer support for interpreting family histories of breast and ovarian cancer in primary care: comparative study with simulated cases. Brit Med J 2000;321:28-32.]

[17] Balter KA, Balter O, Fondell E, Lagerros YT. Web-based and mail questionnaiers: A comparison of response rates and compliance. Epidemiol 2005;16:577-9.

[18] Green MJ, Peterson SK, Baker MW, Harper GR, et al. Effect of a computer-based decision aid on knowledge, perceptions, and intentions about genetic testing for breast cancer suscpetibility. A randomised controlled trial. J Am Med Aassoc 2004:292:442-52.

Address for correspondence

Dr Vitali Sintchenko,
Centre for Health Informatics,
University of NSW
UNSW Sydney 2052 Australia
v.sintchenko@unsw.edu.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Modeling and Acquisition of Drug-Drug Interaction Knowledge

Frédéric Mille[a,b], Patrice Degoulet[a], Marie-Christine Jaulent[a]

[a] INSERM,UMR°S 872, Eq 20, Les Cordeliers, Paris, F-75006 France ; Université Pierre et Marie Curie-Paris6, UMR°S 872, Paris, F-75006 France ; Université René Descartes, UMR°S 872, Paris, F-75006 France;
[b] AP – HP, Hôpital Robert Debré, Pharmacie, Paris, 75019 France, France

Abstract

Objectives*: The effectiveness of computerized clinical decision support systems (CDSS) depends on the quality of the knowledge they refer to. In this article, we are interested in the acquisition, modeling and representation of the knowledge embedded in the "national reference framework of drug-drug interaction" published by the French Health Products Safety Agency.*

Methods*: A model of drug-drug interactions has been designed using bottom-up and top-down approaches. This model is the basis for the design of an XML format to represent and extract information on drug interactions from the reference framework.*

Results*: A specific tool has been developed to extract the information from a corpus of 1053 drug monographs using a methodology similar to the one used by the GEM-Cutter tool to extract information from clinical guidelines. Strategies to integrate the XML files produced into CDDSSs are discussed.*

Discussion-conclusion: *Modeling and acquisition of drug-drug interaction knowledge from a corpus of drug monographs is a potential approach to foster the development of CDSS and improve their specificity.*

Keywords:

drug interactions, knowledge modeling, XML

Introduction

Technology–based interventions have been recommended for reducing the likelihood of medication errors. Computerized physician order entry (CPOE) has been cited as one of the most effective ways to avoid medication errors[1-3] caused by misinterpretation of handwritten orders, incorrect doses, wrong dose forms and inappropriate administration times. The next step is to integrate effective Clinical Decision Support Systems (CDSS) into CPOE systems.

CDSS are information systems designed to improve clinical decision making, for example detecting drug-drug interactions (DDIs) during the order entry process. Indeed, DDI detection systems were among the first CDSS developed. When a potential problem is identified, these systems can provide real-time alerts, allowing the clinician to make appropriate changes before the prescription is finalized.

The effectiveness of CDSS depends on the quality of the knowledge base which they use. Obviously, for detection of DDIs, any CDSS requires knowledge of these DDIs. The main issue addressed here is the acquisition, modeling and representation of the knowledge in this field. In France, the main source of knowledge for DDIs is the "national reference framework of the drug-drug interaction" published by French Health Products Safety Agency "Agence Française de Sécurité Sanitaire des Produits de Santé" (AFSSAPS) [4]. DDIs are described in free text in these monographs and are not directly accessible by a computerized system.

A simplest way to represent knowledge in monographs would be to build manually two-column tables describing the interactions between pairs of drugs. This approach is difficult to maintain as knowledge progresses. The most advanced approach would be automatic free text understanding by Natural Language Processing (NLP) techniques. These techniques could automatically identify concepts in DDI monographs and semantic relations between these concepts. However, the NLP approach is complex and beyond the scope of this work. Between these two extremes, an affordable and plausible approach is to propose tools to structure the knowledge contained in the monographs so as to make this knowledge available.

The aim of the study is to acquire DDI knowledge in a structured knowledge base that could be used by a computerized system. The paper presents a method and a tool for building a knowledge resource usable in the context of prescription systems.

Our work is based on the hypothesis that DDI knowledge formulation has similarities with clinical guidelines knowledge formulation. Indeed, DDI monographs contain recommendations and risk factors that can be considered in the same way as guidelines. Clinical guidelines have become an important medium for the standardization and dissemination of medical knowledge. The "document-centric approach" has been introduced to facilitate the use of guideline knowledge. In the document-centric approach, the original text of the guidelines is systematically marked-up with respect to the model and kept in the form of a structured document.

Selected for best paper award.

The best known instance of "document-centric approach" is the Guideline Elements Model (GEM) which is an XML (Extensible Markup Language) framework based on a hierarchy of concepts describing guideline contents [5].

The encoding of a clinical guidelines using the GEM framework consists of structuring the textual document using the set of XML markups provided by the framework.

The encoding of a monograph consists of structuring the textual document using the set of XML markups. This can be a complex process, as it requires in-depth analysis of the text content. Similarly, for clinical guidelines, substantial variation is observed in the encoding by different users [6]. This would suggest, as concerns our study, that the complexity of manual analysis can affect the structure of the encoded monographs.

To tackle these problems and support the process of manual marking up of monographs, we developed a specific interface, inspired by the GEM-Cutter application [5].

We propose to use the same methodology to structure the information within monographs.

The first step was to identify the relevant information in the monographs so as to build an XML schema related to DDI information. This schema is implemented in a specific environment to acquire descriptions of DDIs directly from text. The second part of our work was to use and evaluate this tool to develop a knowledge base concerning DDI information from 1000 monographs. The practical result is the knowledge base itself. We present the advantages and limitations of our approach and consider the exploitation of the resulting knowledge base.

Materials and methods

Methods

We used a two-step approach to identify the information elements contained in the "national reference framework of the drug-drug interaction" published by AFSSAPS. First, we identified the main concepts of DDI. Second, we used natural language processing tools to analyze the content of DDI texts.

Then we built a XML schema of DDI information, which was evaluated for its ability to represent the initial text information and maintain its meaning.

Finally, we built specific software to write XML files.

Top-down approach knowledge identification

The identification of large classes relied on the manual study of structure in the monographs. It was done based on the reference frame of AFSSAPS.

Bottom-up approach knowledge identification

We applied a data-driven knowledge approach to identify the concepts of the domain present in the monographs. We used a syntax analyzer module based on the hypothesis of similar dependencies between terms which have similar meanings. Then, a pharmacist reviewed the different terms on the basis of shared syntactical contexts and exploited all the network data to cluster and organize the terms.

The lexicon built with this software made possible a second manual step of semantic analysis registering the principal semantic structures contained in the DDI texts, and their frequencies.

Establishing a XML schema

The creation of an XML language to describe monographs requires conceptual modeling of DDIs. We built an XML schema of DDI knowledge from existing knowledge and terminological analysis. This hierarchy of concepts constitutes the skeleton of the XML schema.

Development of an XML Editor

We built X-DIE (XML Drug Interaction Editor), an XML editor with added functionalities to facilitate the editing of XML. X-DIE is a graphic editor which hides the code in the background and presents the content to the user in a more readable format, similar to the version which must be ultimately posted. The interface is based on a multiple windowing system that displays the original monograph, a XML browser and the corresponding XML-encoded text.

Material

The French drug database Theriaque® developed by the Centre National Hospitalier d'Information sur le Medicament (CNHIM) is responsible for the dissemination of independent information about all drugs available in France [7]. The database contains complete information about the drugs: the pharmaco-therapeutic group, the active component, the excipient, the commercial presentation, indications, contra-indications and AFSSAPS reference concerning DDIs. The set of all 1053 monographs of DDIs available in the Theriaque® database constitutes the initial corpus.

The syntax analyser module is SYNTEX, a natural language processing tool that is widely used to build ontologies from texts. SYNTEX performs a syntactic analysis of the sentences of the corpus, and yields a network of the dependence of words and phrases [8, 9].

The PROTEGE editor was used to build and browse the hierarchy of concepts.

The XML schema was built using an open-source XML file editor, JAXE.

The interface for the marking up of monographs was developed in JAVA with the API DOM (Application Programming interface Document Object Model) dedicated to the production and management of XML files.

Results

Knowledge acquisition results

Top-down approach

In the frame of reference of AFSSAPS, a DDI is defined by a pair of protagonists "A + B" who can be an active substance, indicated by their international non-proprietary names, or a therapeutic class, itself being the subject of

interactions "of class". The wording of a DDI is structured in 4 paragraphs: ❶ nature of the risk, ❷ brief mechanism of action (if known), ❸ action to be taken and ❹ degree of constraint.

For the last type of paragraphs four degrees of constraint are possible: ❶ Contra-indication: This is an absolute character and should not be transgressed, ❷ Not advised: The association should generally be avoided, except if rigorous examination of the risk/ benefit ratio suggests otherwise, and imposes close monitoring of the patient, ❸ Cautioned: The association is possible if, in particular at the beginning of treatment, simple recommendations are respected making it possible to avoid undesirable DDI (adaptation of doses, reinforcement of monitoring, be it clinical, biological, ECG, or other), ❹ Take into account: The risk of DDI exists, generally in the form of accumulation of adverse effects; no practical recommendation can be proposed. It is left to the physician to evaluate the appropriateness of association.

In case of a DDI being classified as a contra-indication or not advised, the action to be taken is generally summarized only as a constraint. Cautioned DDI are often associated with recommendations that are simple to implement to avoid the interaction (adaptation of doses, biological controls, etc.). The classification "take into account" is not associated with any practical recommendation because it announces the addition of adverse effects that can only be avoided by using other therapeutic strategies.

This first stage of the analysis generated various categories of knowledge found in the DDI monographs. These categories are considered to be the main concepts to be identified in DDI texts and are used to group terms in the following analysis.

Several classes are distinguished but there is no organization of knowledge in each class.

The level of organization of the knowledge was insufficient for application in CDSS.

Bottom-up approach

The initial corpus contained about 87,707 words. Once processed with SYNTEX, the corpus gave 9393 candidate terms appearing at least twice. The expert selected 2657 candidate terms : ❶ 1349 noun phrases out of 3150, ❷ 1071 nouns out of 3775, 3) ❸ and 237 adjectives out of 336.

These terms have been classified into 5 categories: ❶ pharmacology and pharmacokinetics (931 terms), ❷ galenical, active principle, dose and regimen (936 terms, this category includes the various characteristics of the associations causing DDIs), ❸ physiopathology (282 terms, this group includes all diseases or symptoms secondary to DDIs), ❹ physiology (188 terms), ❺ others (320 terms).

These selected candidate terms were normalized within each category and this gave a final total of 888 concepts.

The combination of bottom-up and top-down approaches enabled us to build a DDI model with 6 main classes (figure 1): risk association, consequence, mechanism, limitation, precaution for use, risk factors.

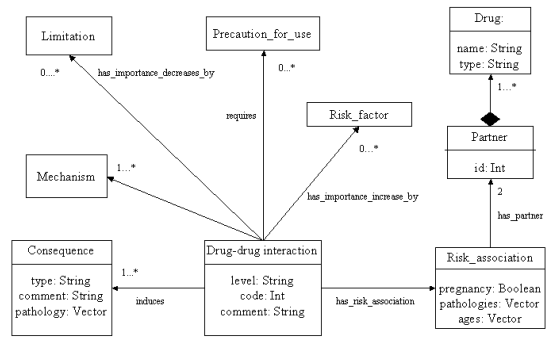


Figure 1 - DDI model

The XML representation schema

The JAXE software allows an XML schema to be built, compatible with the W3C norm.

XML files are constructed as a hierarchy of 72 discrete tags with 6 major branches intended to capture the information in the 6 main sections of a monograph. The 6 branches are:

- *Risk association*: the sub-tree describes the drug association responsible for the DDI, and the clinical situations (pathologies, ages) in which the DDI occurs.
- *Mechanism*: the sub-tree describes the DDI mechanisms (pharmacology, pharmacokinetics, etc.).
- *Risk factors*: the sub-tree describes any physiology (age, pregnancy) and/or diseases that may increase the risk of adverse effects of a given DDI.
- *Consequences*: the sub-tree describes the diseases and symptoms resulting from the DDI.
- *Precaution for use*: the sub-tree describes the precautions to take to reduce the effects of the DDI (substitution, alternative drugs) and/or to reduce its consequences (clinical or biological exams).
- *Limitations*: the sub-tree describes any disease or condition where the benefit for the patient outweighs the risk of an adverse event due to the DDI.

Other types of elements are represented in the XML schema. They include: drug, pathology, biological examinations.

XML elements can have attributes in the start tag. Attributes provide additional information about elements. The attributes of our elements are found in published classifications particularly ATC (Anatomical Therapeutic Chemical), ICD10 (International Classification of Disease – tenth edition) and MedDRA (Medical Dictionary for Regulatory Activities). With these attributes, DDI detection system (DIS) will be able to perform various types of inference between DDI information and the clinical context of the patient. In addition, the XML schema includes

elements to allow quantitative information (for example duration, dose, etc.) to be coded.

The X-DIE environment for the encoding of monographs

The X-DIE environment has been developed to encode monographs. The main interface is a three-windowing system (figure 2), inspired by the GEM-Cutter interface: ❶ window A displays the DDI monograph in full text, ❷ window B displays the hierarchy of the elements for the XML file under construction and ❸ window C displays the information to provide for each element of the file.

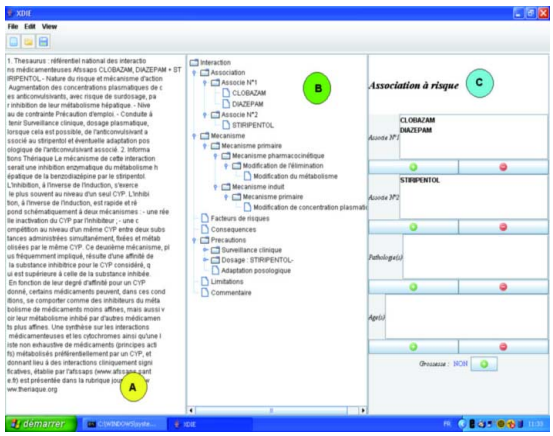


Figure 2 - X-DIE

Our environment allows the user to look for drugs (active principles, therapeutic class and chemical class) for which there is a link, in the data base Theriaque®, with the displayed monograph. It also allows the user to search classifications including ICD10 and MedDra for pathologies similar to those cited in the visualized monograph. Other information is otained by the user.

Knowledge base

One thousand and six of the total of 1053 monographs present in the Theriaque® database could be successfully entered: the coverage was thus 96%. The monographs which were not successfully encoded had at least one of the following two characteristics: ❶ The monograph describes an DDI involving a drug which is no longer marketed; in these cases we chose not to enter the corresponding XML file. ❷ The monograph describes a DDI whose mechanism is a modification of plasma protein binding. This mechanism was omitted by the expert.

Discussion and conclusion

We developed a DDI model capable of representing full DDI information and we developed the X-DIE environment to capture this structured data from full text monographs. We used both general knowledge about DDI contained in the "national reference framework of the drug-drug interaction" published by AFSSAPS, and knowledge extraction with NLP techniques from the same database.

General knowledge helped us to define the generic structure of the DDI model. This top-down approach, although performed manually, rapidly gave an organization of the main concepts, because the domain is delimited. To refine this generic structure, we extracted the knowledge present in the "national reference framework of the drug-drug interaction". For this bottom-up approach, we combined natural language processing results using manual semantic analysis of the relevant DDI candidate terms specifying generic concepts identified by the top-down analysis. The bottom-up and top-down approaches are standard methods in knowledge engineering.

Their implementation in this way ensures that the final structure covers a significant proportion of the information contained in the DDI monographs. However, the percentage of coverage of all the monographs has not yet been established. Moreover, the use of monographs from only one knowledge source potentially imposes limitations: the interoperability of the XML files is reduced in theory because the source for the knowledge and the data to encode is the same. An improvement would be to design the model from different sources of DDI data.

It became clear, during the encoding process, that the monographs were poor in recommendations. A few XML markups were rarely used, for example "clinical examination". We also noted that some recommendations are vague; biological examinations are cited under the generic term "biological monitoring" and clinic examinations are quoted under the generic term "clinical monitoring". This was a surprising finding because monographs may contain: ❶ monitoring options (blood tests, clinical and extra clinical examinations), ❷ guidelines to manage DDIs (modification of the times or sequence of administration, routes of administration, compensation by administration of a third compound or substitution of depleted endogenous substances, adjusting doses, discontinuation as symptoms appear, and suggestions of non interacting alternatives), ❸ factors increasing the risks of DDI (age, sex, genetics, predisposing diseases, and use of alcohol and tobacco).

This work clearly illustrated the paucity of the content of some of the monographs as concerns recommendations. Thus, our model could be used as a tool to enrich and standardize the DDI monographs.

X-DIE tool facilitates the encoding of monographs. However, the data capturing is still largely manual. The encoding process could be considerably simplified if X–DIE could incorporate automatic text processing functions, such as the identification of linguistic markers. Previous work by our group on guidelines demonstrates the value of the identification of linguistic markers [10]. Such extension will exploit the already built ontology to highlight he pertinent information in the text.

Some studies have indicated that the weak specificity and irrelevance of alerts in CDSS lead to loss of confidence in

these systems, a phenomenon known as the "Cry wolf syndrome" [11, 12]. Perceived poor specificity of drug alerts may be a major obstacle to efficient utilization of information and may prevent such alerts contributing fully to improving safety. If there is a false alert, these systems substantially increase the time required to carry out a task, due to the disruption in workflow. High signal-to-noise ratios may also produce alert fatigue and result in physicians skipping past alerts without considering or even reading them. One major limitation of existing DDI detection systems is associated with the use of static knowledge, most often embedded in fixed mapping tables: they proceed by comparing between drugs according to a hard-coded interaction mapping table, without any consideration of the clinical context of the patient or of all the information contained in monographs. These systems are often highly inclusive, placing more emphasis on breadth of coverage than on clinical relevancy or severity of adverse events. In published studies, practitioners explained that CDSS should integrate context relevance information, guidelines or evidence-base medicine. In fact, the majority of DDI can be compensated by dose adjustment or prevented by a well-considered sequence of administration and represents therefore a manageable risk [13]. Physicians wish informative support on DDI, concerning management. A distinction between clinically relevant and negligible DDI is essential [14].

For all these reasons we suggest that the reference framework should be integrated in the form of XML files in CDSS. The current XML-Schema has to be augmented in order to include all information for each DDI in the reference framework: some instructions for laboratory testing to monitor side-effects, replacement of ordered drug with another drug, change in dose, or additional drug. It also contains clinical situations where the use of an association is acceptable in spite of the potential consequences of the DDI; in these clinical situations, it is not necessary to announce the DDI or post an alert. Indeed, in these situations, the clinician can deviate from recommendations for good clinical reasons and the benefits of the drug association outweigh the disadvantages of the potential DDI.

To construct a more accurate and evidence-based CDSS for DDI detection, knowledge from DDI databases as provided by AFSSAPS must be arranged and integrated, but it is also essential to construct algorithms that can exploit this knowledge base appropriately. The next steps include the development of algorithms able to use XML files. Also, the impact of this combination (algorithms and knowledge base) on CDSS specificity will have to be evaluated.

References

[1] Oren E, Shaffer E, and Guglielmo B. Impact of emerging technologies on medication errors and adverse drug events. *American Journal of Hospital Pharmacy*. 2003; 60: 1447-1458.

[2] Bates D, Leape L, Cullen D, et al. Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. *JAMA*. 1998; 280(15): 1311-1316.

[3] Teich J, Merchia P, Schmiz J, Kuperman G, Spurr C, and Bates D. Effects of computerized physician order entry on prescribing practices. *Arch intern med*. 2000; 160: 2741-2747.

[4] Thésaurus des interactions médicamenteuses. 2006, Agence française de sécurité sanitaire des produits de santé: Paris. p.165.

[5] Shiffman R, Karras B, Agrawal A, and Chen R. GEM : A proposal for a more comprehensive guideline document model using XML. *J Am Med Inform Assoc*. 2000; 7(5): 488-498.

[6] Karras B T, Nath S D, and Shiffman R N. A preliminary evaluation of guideline content mark-up using GEM--an XML guideline elements model. *Proc AMIA Symp. 2000*: 413-7.

[7] Amiel M, and Husson M. The THERIAQUE database and information on side effects of drugs. *Therapie.* 1994; 49(5): 455-458.

[8] Fabre C, and Bourigault D. Linguistic clues for corpus-based acquisition of lexical dependencies. *in Corpus Linguistics 2001 Conference. 2001*: UCREL Technical Papers.

[9] Charlet J, Bachimont B, and Jaulent M C. Building medical ontologies by terminology extraction from texts: An experiment for the intensive care units. *Comput Biol Med*. 2005; 36(7-8): 857-70.

[10] Georg G, Colombet I, and Jaulent M C. Structuring Clinical Guidelines through the Recognition of Deontic Operators. *Stud Health Technol Inform*. 2005; 116: 151-6.

[11] Glassman P, Simon B, Belperio P, and Lanto A. Improving recognition of drug interactions : benefits and barriers to using automated alerts. *Medical care*. 2002; 40(12): 1161-1171.

[12] Magnus D, Rodgers S, and Avery A. GPs' views on computerized drug interaction alerts : questionnaire survey. *Journal of clinical pharmacy and therapeutics*. 2002; 27(2): 377-382.

[13] Bergk V, Gasse C, Rothenbacher D, Loew M, Brenner H, and Haefeli W. Drug interactions in primary care: impact of a new algorithm on risk determination. *Clin Pharmacol Ther.* 2004; 76: 85-96.

[14] Bergk V, Gasse C, Schnell R, and Haefeli W. Requirements for a successful implementation of drug interaction information systems in general practice: results of a questionnaire survey in Germany. *Eur J Clin Pharmacol*. 2004; 60(8): 595-602.

Address for correspondence

Dr. Frédéric Mille. UMR_S 872, 15 rue de l'école de médecine, 75006 Paris, France. frederic.mille@rdb.aphp.fr

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

A Systems Development Life Cycle Approach to Patient Journey Modeling Projects

Joanne M. Curry[a], Carolyn McGregor[a], Sally Tracy[b],

[a] *Health Informatics Research, University of Western Sydney, Australia*
[b] *AIHW National Perinatal Statistics Unit, University of New South Wales, Sydney, Australia*

Abstract

Patient Journey Modeling, a relatively recent innovation in healthcare quality improvement, models the patient's movement through a Health Care Organisation (HCO) by viewing it from a patient centric perspective. A Systems Development Life Cycle (SDLC) provides a standard project management framework that can improve the quality of information systems. The concept of following a consistent project management framework to boost quality outcomes can be applied equally to healthcare improvement. This paper describes a SDLC designed specifically for the health care domain and in particular patient journey modeling projects. It goes on to suggest that such a framework can be used to compliment the dominant healthcare improvement method, the Model for Improvement. The key contribution of this paper is the introduction of a project management framework in the form of an SDLC that can be used by non-professional computer developers (ie: health care staff), to improve the consistency and quality of outcomes for patient journey redesign projects. Experiences of applying the SDLC in a midwife-led primary-care maternity services environment are discussed. The project team found the steps logical and easy to follow and produced demonstrable improvement results along with ongoing goal-focused action plans.

Keywords:

medical informatics, SDLC, health care Improvement, patient journey modeling

Introduction

Systems development life cycles (SDLC) were developed to provide a formal structure for the development of quality information systems. An overarching concept of a SDLC is the inclusion of a project management framework for planning, managing and controlling the people, development process and problem solution from the projects inception to the delivery of the required system [1, 2]. This theoretical construct can be applied in a similar manner in a variety of domains including health care improvement.

For a development project to be successful, the people involved in the project must have a detailed plan to follow. Attainment of the required goals depends heavily on having a plan that includes an organized, methodical sequence of tasks and activities that culminate in the delivery of a system that meets the clients' needs for reliability and efficiency. This is a specific goal of a SDLC.

Such concepts can also be aligned to the goals of the dominant health care improvement method, the Model for Improvement (MFI) [3] and it is suggested that by integrating the two approaches, patient journey modeling projects can be conducted in a more consistent manner, delivering higher quality process improvements. In this paper we propose a SDLC approach to Patient Journey Modeling projects that compliments the Model for Improvement via the introduction of a project management framework. Key benefits of the proposed SDLC approach are the provision of a planning, monitoring and control structure that can be used by both IT and non-IT staff with little or no previous process improvement experience to improve the consistency and quality of outcomes for patient journey redesign initiatives.

The paper begins with a background on patient journey modeling and systems development life cycles. Methods and research motivations are then presented, followed by a discussion on the proposed SDLC for patient journey modeling. The paper finishes with highlights of the SDLCs application and a discussion on how the SDLC integrates with the Model for Improvement.

Background

Patient journey modeling

Patient journey modeling is a patient-centric activity that details a patient's progress through a healthcare system for a given service [4, 5]. The goal of Patient Journey Modeling (PJM) is to improve health care quality by reducing variability in the care process. Specifically this includes evidence-based best practice, collecting required information only once, reducing the number of times a patient is moved, eliminating excessive activities, reducing duplicate communications and providing clear and concise information to the patient.

Several terms are used in the literature to refer to the concept of patient journey modeling (PJM) including: clinical pathways, patient flow redesign, clinical practice improvement and redesigning healthcare [6-8].

The most prominent method being used to reengineer health care processes, the 'Model for Improvement' (MFI), provides a framework for developing, testing and implementing changes that lead to improvement. As the method of choice for the Institute of Healthcare Improvement, it has been used extensively in the US [3] and by the NHS in England [4, 5].

Systems development life cycle

A Systems Development Life Cycle (SDLC) is a project management framework that organizes activities into phases [1]. If problem-solving activities are to be productive, the work conducted must be structured and goal-oriented. Computing professionals achieve these results by organizing the work into projects. A *project* is a planned activity (or set of activities) that has a definite beginning and end and that produces a desired result. *Project Management* deals with the planning, monitoring and control of all aspects of a project including the people involved, the problem solution and development process itself.

An SDLC was first introduced to the computing field in the 1960's with the goal of providing guidelines to improve the quality of computer developments [1, 2].

Many systems being developed today follow a development path consisting of 3 core elements: *Analysis, Design* and *Implementation*. *Analysis* activities provide an understanding of the business information system requirements. *Design* activities define the technical architecture and structure of the new systems to satisfy the business requirements. *Implementation* is the actual construction, testing and installation of a functioning information system.

These 3 phases address the core activities required to develop an information system but two additional phases are also required when developing quality systems. A *Project Planning/Initiation* phase involves those activities required to initiate, plan and obtain approval for the project. Once the new system is completed and installed, the development team must perform activities to determine if the project satisfied the original business needs or whether the system needs amendment or enhancement. This is known as the *Post Implementation Review* or *Evaluation* phase.

Current issues

Present patient journey modeling approaches lack any type of project management structure for planning the improvement project or monitoring and controlling its progress. The MFI also lies at a level of abstraction above step-by-step procedures and does not adequately address the integration of technology to assist the change process. The introduction of a SDLC to patient journey modeling projects provides a mechanism for overcoming these issues and provides staff, both IT and non-IT, with a logical guide to achieving improvement results.

Methods and research motivation

The research described uses a constructive research process [9] enriched by aspects of a participatory action research environment [10]. This has involved patient journey modeling sessions with management and staff at Ryde Hospital's Maternity department. Ryde provides midwife-led primary care maternity services for identified low-risk women through the public healthcare system [11, 12].

The work described in this paper originated as part of a Quality Review conducted at Ryde Hospital in 2006. Preliminary analysis of the areas under review indicated that although patient satisfaction was consistently high there were some significant patient assessment, information duplication and system administration issues. An SDLC was designed specifically for the PJM project. This included step-by-step activities and expected deliverables.

A Systems Development Life Cycle for patient journey modeling projects

The proposed Systems Development Life Cycle for patient journey modeling as shown in ***figure 1***, follows the standard SDLC format.

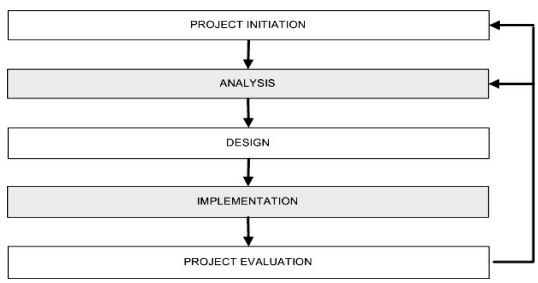


Figure 1 - A systems development life cycle for patient journey modeling

There are 5 phases, namely: Project Initiation, Analysis, Design, Implementation and Project Evaluation. Outputs from the Project Evaluation phase are fed back into the Life Cycle and either trigger new process improvement projects or lead to further enhancements of the resulting patient journey. Figures 2-6 and their description, outline each of the phases in more detail.

Typically senior management will assign a process improvement team leader. In SDLC terms this person is known as the project manager. This role is responsible for reporting progress, leading the team and ensuring that goals and deadlines are met.

Project initiation phase

The primary purpose of the *Project Initiation* phase is to set the scope of the patient journey modeling project and to inform those who will be affected by the results what will occur during the project. ***Figure 2*** shows the 6 major activities that make up this phase with all outputs being stored in a Project Repository.

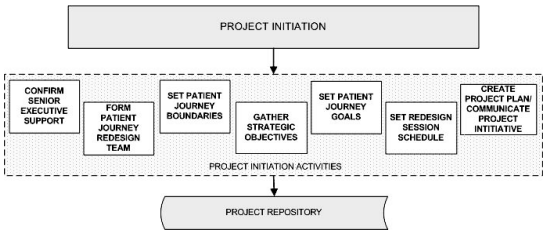


Figure 2 - Project initiation phase activities

The primary contributor to process improvement success is active, strong and visible executive sponsorship throughout the project [13, 14], thus the first activity must be to *secure a sponsor at senior executive level*. This will involve identifying a high-level project 'champion' who will 'talk-up' the project, is authorized to make resource decisions and can report project results to the executive team. Failure to secure a senior executive sponsor may see the project without required resources (both staff and physical) or lead to the project being cancelled if difficulties arise or budgets are tightened.

Once an executive sponsor is confirmed, the project manager sets about *staffing the project team*. In healthcare the inclusion of a representative from all areas affected by the results must be attempted. This includes clinicians, management, IT and administrative staff, and in the case of PJM projects, patient representatives as well. This helps to create a sense of ownership of resulting changes and promotes a culture of on-going process improvement. Ideally staff will be allocated to the project full-time but typically resources can only be released on a part-time basis.

Following formation of the project team the *scope of the patient journey modeling project* is discussed and agreed. The scope sets out what areas are to be included in the project analysis and what areas will be explicitly excluded. This information forms the basis for the first formal documentation created by the project team. It should be recorded in a retrievable medium and stored in the *project repository.* A *project repository* is a central storage area for all project information. Ideally this should be in electronic form as this allows future project teams to easily review past projects, identify successful attributes and reuse them where possible.

Following setting of the project boundaries, strategic objectives relating to the patient journey are gathered. *Strategic Objectives* are set by the Executive and are high level goals and measurements that drive the organisation's overall direction. These may not be clearly defined in some organizations but must be clarified before the project can continue. Specific *patient journey redesign goals* are then defined based on the strategic objectives. Each goal should address a particular area of the problem domain and have expected measurements assigned. These measurements relate to the redesigned patient journey and will determine the degree of improvement attained, post implementation.

Once the team knows what it is trying to improve and how changes will be measured, a *schedule for the running of the redesign sessions* can be created. This aligns with the final activity in the Project Initiation phase, creating the Project Plan. The *Project Plan* lists all of the activities and tasks that will be carried out during the project, estimates their duration and assigns them to a project team member. This plan then guides the progress of the project determining what activities will be conducted, when they must be completed by and who is responsible for their completion. The project manager is responsible for monitoring and controlling the plan and identifying problems. Key information regarding the project initiation phase and the future plan should then be communicated to all areas. This can be done via internal news distributions or as part of in-service sessions. Areas that will be directly affected by resulting changes must be continually kept 'in-the-loop' so that they develop a sense of ownership and there are no surprises during implementation. All documentation created during the phase is added to the project repository for use in the next phase, *Analysis*.

Analysis phase

The *Analysis phase* focus' on creating a graphical representation of the current situation and analyzing how this journey can be improved from the patient's perspective (***Figure 3***).

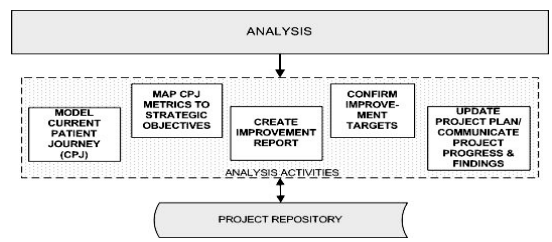


Figure 3 - Analysis phase activities

The first activity is to *create a model of the current patient journey.* This can be done using any process modeling technique including established techniques such as Lean Thinking or emerging techniques such as the 'multi-layered patient flow' communication tool [15, 16].

Existing *measurement criteria* should also be gathered during this exercise. The current patient journey model is created during facilitated group sessions. It is critical during this activity that key stakeholders are involved in the facilitation sessions including administrative, volunteers and patients as appropriate.

Measurements are then *mapped to the strategic objectives* documented in the project initiation phase. This is quite often an enlightening activity, as it will highlight where work is not adding value to the organization, staff or patients.

Following completion of both activities, a comprehensive *Improvement report* is created for management. This details what inefficiencies have been uncovered and what plans of action are possible to improve the current situation. This report should cover areas such as poor use of

human and physical resources, duplicated information collection and communications, unnecessary patient movements, excessive workflow activities and issues of confusion or lack of information for the patient. The report should also recommend priority action areas and activities. The report is submitted to management for discussion and approval.

Once approval is received the *target improvement areas and actions* should be reviewed and confirmed by the project team. It is now time to update the *project plan* based on any new information or scheduling issues and inform the rest of the organization on the outcomes of the analysis phase and what the next steps will be. All documentation created during *Analysis* is stored in the *project repository.*

Design phase

The *Design* phase (***Figure 4***) uses the outputs from the Analysis phase to redesign the patient journey aiming to improve the quality of care and reduce the level of variability for patients experiencing the same journey. It is in this phase that the team will start to work with existing technology constraints and systems and the requirement for new or integrated IT solutions.

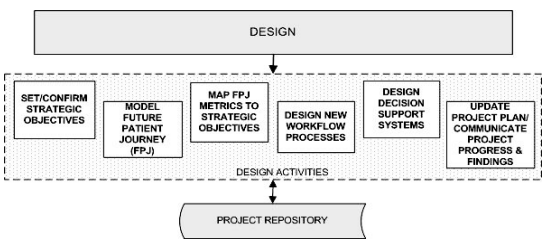


Figure 4 - Design phase activities

The first activity is to *confirm the strategic objectives* that the redesigned patient journey supports.

Once the direction is clear, facilitated group sessions are again used to create a visual representation of the *future patient journey.* This will include new or improved process flows, more efficient use of human and physical resources, streamlined information collection and dissemination, reduced patient movements, improved patient interactions and measurement criteria for all areas.

The *measurement* items are again mapped to the *strategic objectives* to ensure that the new journey is adding value to the organization and its future direction. Some adjustments may be necessary to the measurements defined in the previous activity and as with all other phase documentation these will be updated in the *project repository.*

The completion of the future patient journey and agreed metrics leads to the *design of new or enhanced workflows*. These may be automated or manual workflows. To enable the defined measurements to be captured and analysed, a *decision support system design* is required. This will identify what measurements must be captured and at what stage of the workflow enactment they are required. This is typically the domain of the IT section and will be derived directly from the patient journey redesign work already conducted. The *project plan* is again updated and *project progress and findings are communicated* to the organization.

Implementation phase

The *Implementation* phase (***Figure 5***) is mainly concerned with the development and implementation of the designs output from the Design phase. This will primarily involve the IT section but will still require input from the project team. This will be in the form of system *testing of new/ integrated workflows and the decision support system*.

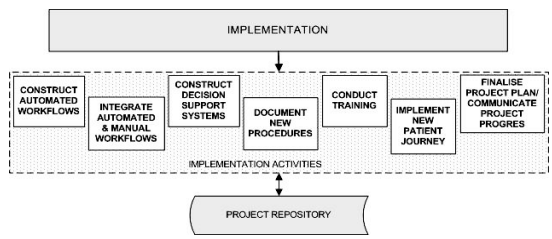


Figure 5 - Implementation phase activities

In parallel to this, the project team will update or *create documentation detailing the new patient journey* along with updated daily workflows. *Training* will need to be conducted on this material as well as any new systems that are to be implemented. The new patient journey goes 'live', with the *implementation of new/enhanced systems and workflows*. The project plan is finalised and details of the project's implementation is communicated. Staff are also advised that following implementation, further improvement is encouraged and can be communicated to the project manager for inclusion in the Evaluation phase.

Project evaluation phase

The *Project Evaluation* phase (***Figure 6***) should be commenced within 3 months of implementation. The main goal of this phase is to revisit the new procedures and determine if they are delivering the expected benefits.

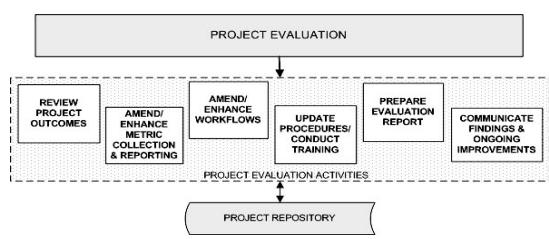


Figure 6 - Project evaluation phase activities

The first activity involves *reviewing the actual results* delivered by the new patient journey and measuring these against expectations. Decision Support System output should be analysed to determine if process metric collection and reporting is meeting set targets *Amendments or enhancements* to the actual measurements and the way they are gathered and reported may be required. This may also lead to *refinement of the newly implemented workflows*.

Any changes to metric collection and reporting or workflows require updating of the documentation in the *project*

repository and possibly republication of updated patient journey procedures. Additional *training* may also be required.

An *evaluation report* is prepared for management outlining evaluation findings and actions taken to further improve the patient journey. Findings may lead to recommendations to revisit the *analysis phase* to conduct a major revision of the (now) current patient journey or may trigger recommendations of completely new patient journey modeling project/s.

Results of the evaluation phase are again communicated to the organization including any new changes or projects.

Results

The Ryde exercise has almost completed the first 3 phases of the proposed patient journey modeling SDLC. Work is still progressing, although some new workflows and procedures have already been tested. This includes the introduction of a new patient assessment form that can be completed and submitted online reducing the number of times a woman is required to attend the hospital prior to her first antenatal appointment. Patient permission paperwork has also been improved reducing the number of forms completed from 3 to 1. Explicit action plans are also in place to complete identified improvements. Health care staff found the SDLC easy to follow and confirmed that when integrated with the MFI, the SDLC had given them a solid direction and set of activities to complete. Specific mention was made of the fact that it could be followed by staff inexperienced in health care redesign projects and made the interactions with the IT section more seamless.

Discussion

Integrating the SDLC with the MFI. This approach is seen as complementary to the Model for Improvement (MFI) not as a replacement. Specific activities within the SDLC align with MFI tasks (ie: set patient journey goals and setting aims) and others can be integrated as part of the MFI (storing gathered information in a project repository). The MFI also gives further information on some of the SDLC activities such as *'Forming the Team', 'Setting Aims'* and *'Defining Measurements'*. An important point to note is how *Plan-Do-Study-Act* (PDSA) cycles are integrated into the SDLC. Once the Design phase is complete, specific areas of the future patient journey model can be selected for implementation on a trial group. Once this trial is complete, the improved procedures can be expanded until management and staff is confident that full-scale implementation should proceed. This means that the Implementation phase can be conducted in an iterative manner, with several iterations leading to implementation of the full future patient journey model over time. The SDLC approach also extends the MFI by including activities for the design and development of technology solutions to support process change. Most importantly the inclusion of a project management framework that supports the planning, management and control of improvement project work provides inexperienced project staff with a solid basis for delivering improvement results within required timeframes.

Conclusion

The integration of a SDLC approach with the MFI for patient journey modeling is achievable and can be understood by all improvement team members, both IT and non-IT. The project team found the SDLC steps logical and easy to follow and produced demonstrable improvement results in the required timeframe, along with ongoing goal-focused action plans.

Acknowledgments

The authors would like to thank the staff of the Ryde Maternity Department at Ryde Hospital, NSW, Australia.

References

[1] Satzinger JW, Jackson RB, Burd SD. Systems Analysis and Design in a Changing World. 3rd ed. Boston: Thomson Course Technology; 2004.
[2] Pressman RS. Software Engineering: A Practitioner's Approach. 4th ed. New York: McGraw-Hill; 1997.
[3] Institute of Healthcare Improvement, The Breakthrough Series: IHI's Collaborative Model for Achieving Breakthrough Improvement. White Paper. Boston: Institute of Healthcare Improvement; 2003, pp. 20.
[4] NHS Modernisation Agency, Improvement Leaders' Guide: Process mapping, analysis and redesign. Best Practice Guidelines. London: Department of Health; April 2005, pp. 22.
[5] NHS Modernisation Agency, Improvement Leaders' Guide: Improving flow. Best Practice Guidelines. London: Department of Health; April 2005, pp. 22.
[6] UK Department of Health, Patient Pathways. Organisation Policy: UK Department of Health; 2005
[7] NSW Department of Health, Easy Guide to Clinical Practice Improvement - A Guide for Healthcare Professionals. Sydney; October 2002, pp. 45.
[8] Szwarcbord M. Redesigning care: improving flow across the hospital [Conference Presentation]. Glenelg: 1st Australasian Redesigning Health Care Summit, Glenelg South Australia, 8-9 March 2005; 2005.
[9] Kasanen E, Lukka K. The Constructive Approach in Management Accounting Research. Journal of Management Accounting Research 1993;5:pp.243-265.
[10] Avison D, Lau F, Neilsen PA, Myers M. Action Research. Communications of the ACM 1999;42(1):pp.94-97.
[11] Tracy S. Proposal for a primary health care model of maternity services at Ryde Hospital - Introduction of the Ryde Midwifery Group Practice. Unpublished Report. Sydney: Midwifery Research & Practice Development Unit; 15th March 2004, pp. 44.
[12] Australian College of Midwives, National Midwifery Guidelines for Consultation and Referral. Guideline. Canberra: Australian College of Midwives; January 2004
[13] Change Management Learning Centre, 2005 Best Practices in Change Management. Benchmark Report. Loveland, CO: Prosci, pp. 51.
[14] Kotter JP. Leading Change. Boston: Harvard Business School Press; 1996.
[15] Institute of Healthcare Improvement, Going Lean in Health Care. Innovation Series Report. Cambridge, MA,: Institute of Healthcare Improvement; 2005, pp. p.24.
[16] Curry JM, McGregor C, Tracy S. A Communication Tool to Improve the Patient Journey Modeling Process. 28th IEEE International Conference of the Engineering in Medicine and Biology Society (EMBC2006); New York, USA.

Address for correspondence

Joanne Curry
email: jm.curry@uws.edu.au)

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

The Nurse—Patient Trajectory Framework

Gregory L. Alexander[a]

[a]*Sinclair School of Nursing, University of Missouri—Columbia, USA*

Abstract

The development of nursing knowledge should give structure and form to the practice of nursing. The development of Nursing Process Theory resulted from early nursing observations and inferences from nursing practice that resulted in formal data accumulation processes, mutual correspondence between nurses and patients, and exchange of information. The development of the nursing process discipline helped to substantiate the need for professional nursing services. The shifts towards examining the links between processes and outcomes, professional accountability, and classification of distinct nursing functions have influenced the development of information systems. The Nurse—Patient Trajectory Framework described in this paper may be used to show the relationships between the virtual information system and the real world that it affects. The framework is visualized along two separate and distinct nurse and patient trajectories.

Keywords:

nursing informatics, information systems, nursing process

Introduction

Nurses and patients interact with computers on many levels. Staff nurses collect information from their clients and, in some cases, simultaneously store client information in information systems. Nursing administration or researchers can use the information contained in this warehouse of data to evaluate utilization, financial impact, and risk management activities of nursing services. The information can also be used to maintain an ongoing record of events, actions, behaviors, perceptions, and progress made during healthcare encounters. Effectiveness of nursing information systems to understand and predict nursing and patient outcomes depends on the usability of the system, organizational workflow, and satisfaction during nursing interactions with computer information systems. Therefore, understanding interactions between nurses, patients, computers, and other elements of information technology has the potential to improve nursing care processes and, subsequently, patient outcomes impacted by nursing services. The following paper describes a framework for nursing informatics (NI) called the Nurse—Patient Trajectory Framework. The framework utilizes nursing process theory, human computer interaction, nursing and patient trajectories as components of a framework that can be used to evaluate patient care systems.

Background

Nursing process theory

Nursing, as a practice discipline, should be concerned with the development of nursing knowledge and nursing knowledge should be used to give structure and form to the practice of nursing [1]. From the roots of the beginning of professional nursing education the nursing process was taught as a means to structure nursing care. Original components of Nursing Process Theory were developed through extensive clinical observations and evaluations of nurses performing the nursing process [2-3]. Emphasis on these clinical observations and inferences made from nursing practice was seen as a forerunner to the nursing process, with its respective components being precursors to a formal data accumulation process [3]. In early anecdotal observations emphasis was placed on the importance of reciprocal relationships between patients and nurses and the importance of the process of nursing care. The nursing process represented the first attempts to develop reciprocity between patients and staff. Reciprocity was garnered through mutual correspondence between patient and staff, through mutual exchange of privilege, and through the mutual dependence, action and influence the patient and staff have on one another.

The four practices basic to the nursing process, as recognized by Orlando [2], were observation, actions, reporting, and recording (Figure 1). Observations included direct or indirect information obtained about a patient while on duty. Direct information was derived from the nurse's own experience of patient behavior. Indirect information included reports of actions, records, or reports of other nurses or allied health professionals. Actions, such as the ability to make decisions or planning care, occur within a context and environment. These contexts are highly influenced by organizational design, area of application, characteristics of the decision maker, maturity of the setting, and importance of the decision [4]. Finally, recording and reporting of nursing information regarding observations and actions is a fundamental function of nursing practice. The effectiveness of nursing process is dependent upon the clinical inferences made from information captured. Every clinical inference made involves an element of risk for the nurse, patient, and the relationship between them [3].

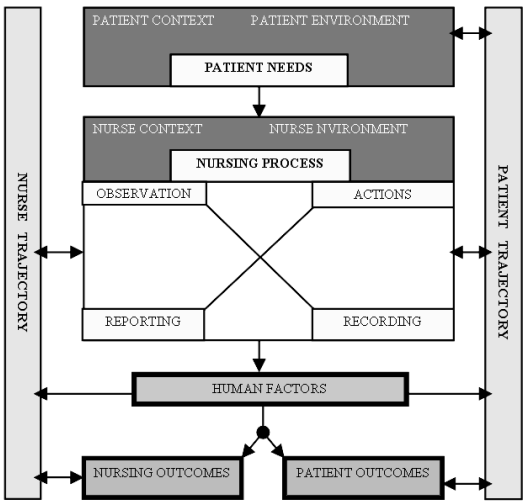


Figure 1 - The Nurse-Patient Trajectory Framework

Nursing informatics: creating the interface between nursing process and technology

Werley and Grier, two pioneers in nursing informatics, advocated for the development of nursing information systems(IS)[5]. Their work led to the identification of nursing data elements essential to diagnosing nursing problems, choosing nursing actions, and evaluating the nursing process in electronic health records (EHR). Furthermore, they suggested research directions to promote the development of technology in caregiving activities. Finally, Werley and Grier published one of the first nursing models that established a link between IS and the development of nursing knowledge.

Their model integrated community data, institutional data, interpersonal data, and patient data into a hierarchical framework. It was suggested that these information sets were needed to assist nurses in making decisions at various levels of functioning[6]. Sets of information thought to be important to nursing decisions included aggregate data on communities including population risk, cultural status, institutional data on finances and structure of facilities, interpersonal elements of caregiving including nursing interventions, orders, and outcomes, and finally, patient data including diagnoses, psychosocial factors, and patients perceptions and goals. Facets of this systems model approach can be seen in subsequent NI frameworks [7;8].

Schwirian developed the NI pyramid as a model for Nursing Informatics [9]. Schwirian defined NI as, "the use of information technology in relation to any of the functions which are within the purview of nursing and which are carried out by nurses."(p. 134). NI activity was depicted as an interface between the computer hardware and software, raw nursing related information, and the user within the context of their profession or organization. All of these elements led to a common goal or objective. The model is described as being flexible and multidimensional allowing the researcher to enter into the model at various points depending on the research questions and hypotheses posed.

Graves and Corcoran [10] defined NI as a, ". . .combination of computer science, information science, and nursing science designed to assist in the management and processing of nursing data, information and knowledge to support the practice of nursing and delivery of nursing care" (p.15). These authors emphasized the processing of nursing information as it progressed from data to information, and finally, to nursing knowledge. In a related article, Goosen [11] extended Graves and Corcoran's NI model to include decisions made in clinical practice, activities that follow nursing decisions, and in the final evaluation, consideration of patient outcomes. These models emphasized the importance of understanding how nurses utilize information to develop knowledge. Goosen went one step further to include pragmatic aspects of information or how information leads to nursing actions.

Turley [12] described a model for NI based on three themes derived from past definitions of nursing informatics. The themes regarded the use and the position of the computer and computer science in informatics, conceptual issues, and functional performance in informatics. These themes underscored the important role computer technologies play in the daily functions of nurses. The model utilized nursing science as a base of knowledge to promote the advancement of nursing informatics as a discipline. Computer science, information science, and cognitive science were represented as spheres in the model that overlay nursing science. The juncture between the three spheres represented the informatics domain. Computer science represented the development of hardware and software to facilitate new understanding and new ways of representing knowledge. Information science facilitated knowledge of organizational structure and informational flow through the organization. Finally, Turley indicated that cognitive science helps to clarify information technology by improving information retrieval, perception of information encountered, and understanding of information processing. Staggers and Parks [13] developed a model called the Nurse—Computer Interaction Framework which has been used to help understand interactions between nurses, computers, and enabling elements that optimize the ability of nurses to process information via computerized systems. The authors identified five elements commonly included in human—computer interaction (HCI) frameworks including the user, computer, tasks, interfaces, and environmental elements. After a review of several frameworks the authors reached several conclusions regarding previous NI frameworks including: 1) most frameworks lack environmental and task oriented elements that are essential to understanding computer interactions, 2) elements of frameworks are conceptualized differently across different frameworks, and 3) most frameworks do not include a dimension of time.

Staggers and Parks [13] included a developmental trajectory for NI including time dimensions not previously developed in NI models. According to the model the NI trajectory has important implications because: a) nurse—computer inter-

actions can change over time and b) the location of phenomena along the trajectory has important implications for outcomes related to nurse—computer interactions.

The most recent model of NI described by Effken [8], the informatics research organizing model (IRO), extends Donabedian's structure—process—outcomes model and emphasizes elements of Nursings' metaparadigm including the system, nurse, patient, and health. Effken described the IRO model as being highly abstract and as being able to accommodate various middle range theories and conceptual frameworks. Effken indicated that all organizing frameworks for NI must address and represent two essential components including context and components of nursings' metaparadigm. Based upon these criteria, previous frameworks for NI were found to exclude specific elements of these essential components. According to Effken, some of the current NI models do not explicitly make the patient part of the model, while other models do not define the context or include all elements of nursings' metaparadigm.

Nurse-patient trajectory framework

Previous NI models have been criticized for not explicitly including aspects of patient care but, being more about nursing management than patients [8]. As discussed previously, reciprocal relationships between patient—nurse, nurse—nurse, and the nurse—significant other are integral components that need to be included in the clinical decisions made by nurses. Evidence of these relationships is represented in the knowledge gained through interactions between these individuals. These interactions facilitate an exchange of communication between patient and nurse that lead to better understanding of the contextual and environmental factors attributed to each person. Crucial factors that must be recognized in shared information are cultural, social, economic, and physical characteristics; excluding this information interferes with the ability to fully understand potential outcomes of a patient [8]. Including this information can facilitate more effective nursing actions that can lead to better individual outcomes along nurse and patient trajectories.

Defining nurse and patient trajectories

The term trajectory in health care can be defined as the assembling, scheduling, monitoring, and coordinating of all steps necessary to complete the work of patient care. The term trajectory refers not only to the pathophysiological process of a patients disease state, but also refers to the total organization of work done throughout all nurse and patient interactions and refers to the impact of patient care processes on those interactions and the organization[14;15]. Trajectories involve different medical and nursing actions by people with different types of skills and resources, trajectories lead to a separation of tasks between workers, including kinfolk and the patient, and trajectories must consider the different relationships between all workers[15].

Two separate trajectories, the nurse trajectory and patient trajectory, are identified in the proposed framework (Figure 1). While appearing to be in parallel with each other these trajectories could be viewed as quite independent of each other. Associated with each trajectory is a trajectory scheme that can be imagined as a sequence of potential events or anticipated events along the trajectory[15]. The beginnings of the trajectory may have two different dimensions for the nurse or patient. The nurse may characterize the diagnosis or chief complaint as the beginning of the trajectory. A patient's trajectory may begin when a symptom or a need appears before coming in contact with a health care professional.

The patient's context and environmental characteristics are also seen as separate and different from nursing context and nursing environment. Context has been described as a multilayered construct that has cultural, economic, social, and physical implications for understanding potential and actual outcomes[8;13]. These actual and potential outcomes are associated along two trajectories, one for nursing and one for patients. The potential and actual outcomes are affected by how technology is integrated into the environment and by the users ability to interact with technology.

Nurse trajectories

Nurse trajectories begin when the diagnoses or chief complaint is determined. Nursing contexts are described in the observations, actions, reports, and records of nursing information. Patient behaviors and perceptions of the nurse that is described in the context of the nursing data influence clinical decision making. Decisions may be influenced by accessibility of information, how information is classified and stored, how it is communicated, how technology is used, and design of workspace including both physical and virtual environments (Salvendy, 2005). Finally, a set of nursing outcomes is identified on the nursing trajectory. While these nursing outcomes may certainly overlap with patient outcomes (i.e. patient safety) the implications for nursing will be different than for the patient. For example, if a nursing process is changed related to medication administration practices the nurse might require education of new policy changes and possibly competency evaluation while the patient only knows that they have received the right medication, just in time and at the right dose.

In contrast to those who attempt to define the nursing process through more descriptive measures, other evaluations of nursing processes using EHR center on quality of documentation, patient satisfaction, and nurse perceptions. Evidence of the benefits and the lack of benefit of IS that incorporate nursing documentation and case management strategies have been reported [16-18]. Nahm and Poston evaluated an integrated point of care systems effect on nursing documentation. The authors identified several attributes of computerized IS that contribute to quality documentation including: a) prompts or reminders within assessments and interventions to alert nurses to required documentation, b) ability to collect real time nursing data, c) standardized, streamlined assessments and interventions in menus and interfaces, d) mandatory fields requiring nursing attention before the nurse can proceed, e) information retrieval from past visits, and f) incorporated work tools that sequence and consolidate tasks and provide reminders when part of the nursing process is missed [16].

The authors also found that computerized documentation did not interfere with patient satisfaction.

In another study evaluating nursing documentation pre- and post-implementation of IS the authors found that IS did not significantly improve documentation within the first 6 months of the study [17]. However, with re-education of nurses on the use of IS documentation of assessments of outcomes, goals, and nursing interventions performed did improve by the end of the 18 month post-implementation phase[17]. Nurse perceptions of clinical information systems were evaluated to determine different views between computer users and non-users about how IS affected their practice [19]. Interestingly, this study indicated that there was a significant difference between the two groups when asked about satisfaction and professional status with computerization. Nonusers were less satisfied because they felt the computer interrupted their thought processes, they felt they could not trust the computer, and they felt the computer thought to much for them resulting in a reduced professional status[19].

Patient trajectories

Patient trajectories may begin with an identified need or symptom and are dependent on a separate set of contextual and environmental factors than healthcare workers. Patient trajectory schemes may be well developed and thought out before they even have any contact with healthcare providers. At the onset of the identified need or symptom patients may begin accessing healthcare information via the World Wide Web or other sources so that they are armed with information for the healthcare worker by the time the diagnosis or chief complaint is made.

There are cultural, economic, social, and physical considerations within the context of the patient environment that shape the patient trajectory[7]. For example, a patient's physical location can have implications for the availability of medical technology or other sources of health care such as information on the Internet. Patient trajectories may also be influenced by past personal experiences or by relationships with other people with similar needs.

The use of technology to evaluate nurse-patient trajectories

Technology can influence trajectories by producing an entirely new trajectory or by lengthening trajectories[15]. New trajectories are created when medical information, that was previously difficult to find, is found with a simple keystroke, with an automated computer alerting system, or by creating color changes in critical text fields holding vital clinical information. The new information may lead to different clinical decisions or judgments regarding treatments, potential outcomes, diagnoses, or utilization of health care resources.

The lengthening of trajectories creates new medical, organizational, and personal problems for patients who are living longer than expected and thus have more complex illnesses and trajectory schemes. Lengthened trajectories may lead to increased specialization, costs, and oftentimes uncertainty of outcomes. Although, research on nurse and patient trajectories is limited some research exists that can be used to describe trajectory evaluation[20-22].

Nurse-patient perceptions and human-computer interaction

The complexities of organizing therapeutic actions are derived from the multiple trajectories, the range and number of complex tasks, which affect the course of patient care and the organization of those tasks [14-15]. The ability of nurses to perceive and organize their work in IS depends on how nurses interact with the computer systems.

Human computer interaction (HCI), oftentimes used interchangeably with usability or human factors, addresses specific issues of human performance during computer interactions including ease of learning, use, remembrance, satisfaction, efficiency, error-forgiving interactions, and seamlessness of fit to tasks [23]. Previous research has shown that IS can provide benefits by seamlessly linking homebound persons with Alzheimer's disease and caregivers with information resources, online health related support, and training [7]. The information system was found to provide strong interpersonal support [7]. This type of early intervention via a computer linking patient, caregiver, and healthcare provider will allow for earlier intervention by nurses in the patient trajectory scheme and may improve patient outcomes.

In other studies, research has shown that understanding the interface between users and the information system plays a role in nurse-patient trajectories. Healthcare providers are challenged by the availability of information at the point of care. In a study, designed to discover and implement design principles to facilitate healthcare practitioners access to healthcare information, a strong correlation was found between total time, navigational ability, and perceived functionality within a computer interface[24]. Ability to find accurate information on which to base decisions can affect the ability of nurses to provide care that is evidence based. Further studies have shown that by reducing the barriers to the use of clinical reminders, such as usefulness, workflow, and efficiency, quality of care of inpatients may improve [25].

Conclusion

The purpose of a framework or a model is to show varying degrees of relevance between virtual or imagined systems and the real world it represents [1]. The purpose of this paper is to describe a framework that incorporates the fundamental components of nursing, the nursing process, with principles of human computer interaction. The framework can be visualized along two separate and distinct trajectories, nurse or patient, which ultimately, depending on the design of the information system, may impact nursing and patient outcomes.

Acknowledgement

Supported by the National Institutes of Health T-15-LM07089-12 National Library of Medicine Fellowships.

References

[1] Young A, Taylor SG, McLaughlin-Renpenning K. Connections: Nursing Research, Theory, and Practice. St. Louis: Harcourt; 2001.

[2] Orlando I. The Dynamic Nurse Patient Relationship. New York: National League for Nursing; 1961.

[3] Kelly K. Clinical inference in nursing. Nursing Research 1966;15(1):23-6.

[4] Abbott PA, Zytkowski ME. Supporting clinical decision making. In: Englebardt SP, Nelson R, editors. Health Care Informatics: An Interdisciplinary Approach.St. Louis: Mosby; 2002. p. 115-28.

[5] Nursing Information Systems. New York, NY: Springer; 1981.

[6] Newcomb BJ. Issues related to identifying and systematizing data--group discussions. In: Werley HH, Grier MR, editors. Nursing Information Systems.New York, NY: Springer; 1981. p. 278-96.

[7] Brennan PF. Characterizing the use of health care services delivered via computer networks. Journal of the American Medical Informatics Association 1995;2(3):160-8.

[8] Effken E. An organizing framework for nursing informatics research. Computers Informatics Nursing 2003;21(6):316-25.

[9] Schwirian PM. The NI pyramid: A model for research in Nursing Informatics. Computers in Nursing 1986 May;4(3): 134-6.

[10] Graves JR, Corcoran S. The study of nursing informatics. Image Journal of Nursing Scholarship 1989;21(4):227-31.

[11] Goossen WTF. Nursing information management and processing: A framework and definition for systems analysis, design, and evaluation. International Journal of Bio-Medical Computing 1996;40:187-95.

[12] Turley J. Toward a model for Nursing Informatics. Image Journal of Nursing Scholarship 1996;28(4):309-13.

[13] Staggers N, Parks PL. Description and initial applications of the Staggers & Parks Nurse--Computer Interaction Framework. Computers in Nursing 1993;11(6):282-90.

[14] Goorman E., Berg M. Modelling nursing activities: Electronic patient records and their discontents. Nursing Inquiry 2000;7(1):3-9.

[15] Strauss A, Fagerhaugh S, Suczek B, and Wiener C. Social Organization of Work. Chicago, IL: University of Chicago Press; 1985.

[16] Nahm R, Poston I. Measurement of the effects of an integrated point of care computer system on quality of nursing documentation and patient satisfaction. Computers in Nursing 2000;18(5):220-9.

[17] Larrabee JH, Boldreghini S, Elder-Sorrells K, Turner Z, Wender RG, Hart JM, et al. Evaluation of documentation before and after implementation of a nursing information system in an acute care hospital. Computers in Nursing 2001;19(2):56-65.

[18] Williams SU. Computerized documentation of case management: From diagnosis to outcomes. Nursing Case Management 1998;3(6):247-54.

[19] Axford R, Carter B. Impact of clinical information systems on nursing practice: Nurses' perspectives. Computers in Nursing 1996;14(3):156-63.

[20] Hawthorne MH. Using the trajectory framework: Reconceptualizing cardiac illness. Scholarly Inquiry for Nursing Practice 1991;5(3):185-95.

[21] Corbin JM. The Corbin and Strauss Chronic Illness Trajectory model: An update. Scholarly Inquiry for Nursing Practice 1998;12(1):33-41.

[22] Alverzo J. The use of aesthetic knowledge in the management of brain injury patients. Rehabilitation Nursing 2004;29(3):85-9.

[23] Staggers N. Usability Concepts and Clinical Computing. In: Ball MJ, Hannah KJ, Newbold SK, Douglas JV, editors. Nursing Informatics: Where Caring and Technology Meet. 3rd ed. New York: Springer; 2000. p. 95-109.

[24] Alexander GL, Hauser S, Steely K, Ford G, Demner-Fushman D. A Usability Study of the PubMed on Tap User Interface for PDAs. Proceedings MEDINFO 2004. 2004.

[25] Patterson ES, Nguyen AD, Halloran JP, Asch S.M. Human factors barriers to the effective use of ten HIV clinical reminders. Journal of the American Medical Informatics Association 2004;11:50-9.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

System Analysis and Improvement in the Process of Transplant Patient Care

Catherine J Staes[a], R Scott Evans[a,b], Scott P Narus[a], Stanley M Huff[a,b], John B Sorensen[c]

[a]*Department of Biomedical Informatics, University of Utah School of Medicine, Utah, USA*
[b]*Intermountain Healthcare, Utah, USA*
[c]*Department of Surgery, University of Utah School of Medicine, Utah, USA*

Abstract

Clinical information c oncerning transplant patients is voluminous and difficult to manage using paper records. A system analysis was performed to assess information system needs of the liver, kidney, and pancreas transplant program at LDS Hospital in Salt Lake City, Utah. After evaluating workflow, decision support needs, and requirements, we designed and implemented an extendable information system to support care following liver transplantation. We developed and implemented a standardized operative note, forms to enter external laboratory results and transplant-related information into the electronic health record, and computerized alerts to notify the transplant nurses when liver transplant patients had new, abnormal, or overdue laboratory results. The information system has improved the quality of clinical data available in the EHR, clinician satisfaction, and efficiency with management of laboratory results. The components developed for this project can be extended to meet other transplant program needs.

Keywords:

system analysis, ambulatory care, transplantation

Introduction

In the United States during 2005, 28,107 persons underwent solid organ transplantation for end-stage disease.[1] Long-term survival depends on the patient's clinical status, the surgical procedure, the donated organ, and the management of immunosuppressive therapy and complications. The process of evaluating patients, matching them with donors, and monitoring them after transplantation, generates large volumes of information from multiple sources. Physicians, nurses, and support staff that work with transplant programs need to manage and access this large volume of laboratory, clinical and other data to make decisions.

LDS Hospital, in Salt Lake City, Utah, has a liver, kidney, and pancreas transplant program that selects and monitors adult patients that reside in eight western states in the United States. LDS Hospital is 1 of 21 hospitals and numerous outpatient facilities included in an enterprise called Intermountain Healthcare (IHC). In 2004, 159 kidney, liver or pancreas transplant surgeries were performed at LDS Hospital, and 1,216 transplant patients required outpatient monitoring. The volume of information managed by the transplant program increases each year as the annual number of transplantations increases and survival rates improve.[2] In 2000, the director of the transplant program at LDS Hospital requested assistance with developing "a database". Their information system was almost completely based on paper records. The transplant program could purchase or build a stand-alone information system, or they could integrate their needs with electronic record systems available at LDS Hospital.

The objectives of this project were to define the requirements for a transplant program information system, to identify a transplant process that could be improved with computerized information technology, to initiate system development, and to make a positive impact on patient care.

Methods

Using a systematic approach [3], we addressed each software development phase and used our findings to inform the next phase. We performed a system analysis, requirements analysis, and defined a feasible project scope. The system was designed and developed using knowledge engineering methods, vocabulary development, and application programming. We performed usability testing and quality assessment on the new computerized components, implemented the system, trained the transplant team, and evaluated the impact and user satisfaction. Selected methods and results are reported.

System analysis

System analysis is important for understanding problems, directives, and opportunities; establishing priorities; and assessing feasibility.[4] Before expending resources, the following questions needed answers. What information was needed to manage transplant patients? Who needed information and from where will it be accessed? What processes would benefit from computerization? Workflow and information flow were assessed by attending weekly transplant meetings and observing processes in the office and clinic. Existing data forms, reports, and record systems were assessed. The medical, nursing, and support staff were interviewed to determine priorities and clarify the process of transplant patient care. We reviewed clinical

transplant systems described in the literature[4-6] and stand-alone systems available from software vendors.

Application development

The IHC information system had functionality that could meet some transplant program needs.[7] The transplant clinicians had access to the LDS Hospital clinical information system "HELP" and the IHC enterprise-wide, longitudinal electronic health record (EHR).[7] The EHR included laboratory results from 16 IHC hospitals and five IHC clinic laboratories. Within the EHR, clinicians could view IHC laboratory results and enter and view medications, allergies, problems, and clinical notes. A patient list function in the EHR was used to specify patients whose data should be used to trigger alerts and create reports. An application was available to build data entry forms to enter structured laboratory and transplant data. An integrated decision support system was used to create a logic module for transplant related alerts. An integrated messaging application was used by clinicians to view alerts and laboratory data, and navigate within the EHR.

Assessment of data quality on the paper flowchart

An oversized paper flowchart (11 by 25.5 inches) was the primary record used by the transplant team for tracking patients after transplantation. The paper flowchart was maintained and continued to be the primary clinical longitudinal record throughout the entire project. The quality of laboratory data on the paper flowchart is an indicator of information flow and the quality of information being used for decision- making. The completeness of creatinine, tacrolimus, and cyclosporin A results transcribed onto the paper flowcharts was assessed before and after the alerts were implemented. On the day of a patient's chart review, the previous four months of computerized IHC laboratory results for the patient were printed. Each computerized result was compared to the result recorded on the paper flowchart and classified as a "perfect match," "missing," or "different" (incorrect value or specimen collection date). We excluded results that were collected before the transplant program managed the patient. We included only one result each day because only one result is charted on the paper flowchart each day. Both inpatient and outpatient IHC results were included.

Results

System analysis

Information required for transplant patient care was located in the LDS Hospital "HELP" system, the IHC enterprise EHR, and the inpatient and outpatient paper record systems at LDS Hospital. Additional transplant patient information was stored in at least 10 other paper and electronic record systems. . During the study, the flowchart was found to be useful for integrating IHC and external laboratory results in chronological order so clinicians could view trends in laboratory results and immunosuppression drug dose and levels side by side. In addition, the flowchart included patient demographics, interventions, complications, and some nursing documentation. The paper flowchart was not always accessible, was time-consuming to maintain, difficult to reproduce if misplaced, and not amenable to computerized decision support or data analysis.

Six major processes associated with the management of liver, kidney, and pancreas transplant patients were identified (Figure 1). During the final process, clinicians monitored patients after transplantation to manage immunosuppression therapy and to prevent organ rejection, infections, and medication toxicity. Each process involved a unique set of patients, records, and activities; however, common features were identified. All six processes required the reporting of information to the national organization United Network for Organ Sharing (UNOS), and communication with external entities.

The system analysis generated the following findings:

- Among all patients managed by the transplant nurses, more patients are managed after transplantation (77%) than prior to transplantation (23%) when they are being evaluated or waiting for transplant. The population of post-transplant patients is increasing (an average of 4% each year between 2000 and 2004).
- The transplant clinicians closely manage liver transplant patients for their lifetime. Kidney and pancreas patients are transferred to community physicians a few months after transplantation.
- Approximately 70% of the laboratory results for liver transplant patients are available in the EHR. Clinicians report that laboratory results are the most important clinical data used for managing patients after transplantation.
- When managing patients after transplantation, nurses repeatedly performed a task amenable to computerization and decision support: they perform surveillance on laboratory results and observe occurrences, trends, and omissions. In contrast, prior to transplantation, clinicians gathered a heterogeneous set of information to make one decision (e.g., Is the person eligible for transplantation?). Also, prior to transplantation, there is a greater need to exchange information and establish interfaces with outside entities.
- The nurses were concerned about losing patients to follow-up. Transplant patients need periodic testing for immunosuppression drug levels and kidney toxicity to recognize complications and monitor medication tolerance. Laboratory testing is performed three times a week soon after transplantation, then decreased over time to every three months. The nurses requested a system for tracking laboratory results to identify new, critical, and overdue laboratory results.
- Transplant management information systems available from vendors did not meet our needs for several reasons. Interfaces would need to be developed to download laboratory and other clinical data from the EHR. Information charted in the transplant information system would not be accessible for viewing and decision support from the EHR and would be yet another electronic data storage silo. The system would duplicate functionality that already existed in the EHR.

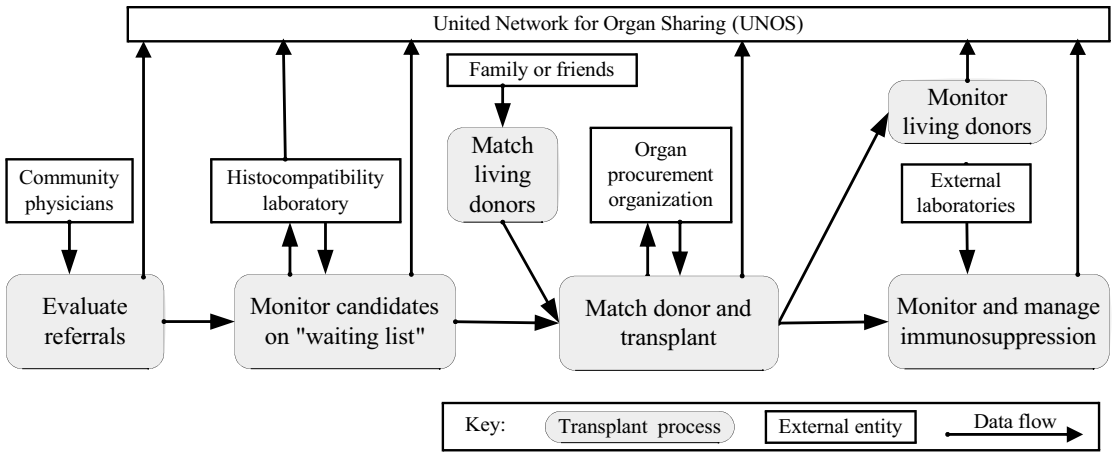


Figure 1 - Major processes associated with the management of solid organ transplant patients

Finally, external laboratory results would still need to be hand entered into the system to integrate external and IHC results.

- Information required for UNOS reporting was not always available in the current free-text operative note. The methods used for calculating ischemia times varied between clinicians.

Project scope

A project targeting the management of patients after liver transplantation was selected for the first system improvement project.

Requirements

The following major requirements were identified:

1. Identify the patient population and each patient's status with the program.
2. Use models that can be extended to other organs and processes of transplant care.
3. Use vocabulary useful for clinical care, but consider concepts reported to UNOS when creating lists of "reasons for transplant" and data collection forms.
4. Improve the quality of documentation about the transplant event, the donor, and risk factors so information is available for decision support, reporting to UNOS, and analysis of outcomes.
5. Integrate external and IHC lab data in the EHR.
6. Create a side-by-side view of laboratory results and medication dosage.
7. Meet regulatory and accreditation standards.
8. Use and do not duplicate functionality that exists in the IHC EHR.

System Design

Improved data collection

A new paper operative note was developed with input from the surgeons. This note standardized documentation about the transplantation and included a graphic to calculate organ ischemia times. This information is important for both clinical management and accurate reporting to UNOS.

Model development and data input

After defining core data fields and improving documentation, data models and entry forms were developed to store information in the EHR. Forms were created to enter the patient's status with the program, selected laboratory results reported by laboratories external to IHC[8], and core information about the transplantation, the donor, and risk factors present at the time of transplantation. External laboratory results in the EHR were critical for automating surveillance of overdue laboratory results and avoiding false overdue alerts. Donor and risk information is important for decision making (e.g., CMV prophylaxis) and program evaluation.

Data views and reports

A "Transplant" view of laboratory results was created so clinicians can see external and IHC results in chronological order with analytes displayed as they are ordered on the paper flowchart.[8] Reports were developed to support workflow and summarize transplant patient information.

Computerized alerts

The transplant physicians and nurses defined logic used to generate alerts. Alerts were received for all new creatinine results (Figure 2). An additional message was included when creatinine results showed an acute or trending increase of 0.3 mg/dL or more. The creatinine alerts prompted the nurses to view all new laboratory results in the EHR. Alerts were also received for all new immunosuppression drug levels and for abnormal potassium and magnesium values. Each alert message included the date of transplantation and the time since transplantation, and indicated if the patient was hospitalized. Alerts were also received when patients were overdue for creatinine or immunosuppression drug level testing. For example,

patients that underwent transplantation during the previous three to six months are expected to get their blood tested every two weeks. If a patient had no creatinine or immunosuppression drug level for 21 days, then an ìoverdueî alert was triggered. When reviewing and acknowledging alerts, the nurses documented their actions in the EHR.

Implementation

In February 2003, the "Transplant" view of laboratory data was implemented. In spring 2004, the surgeons started using the new operative note; the medical assistant started entering external laboratory results; the nurses started entering information about transplantations; the patient list was generated and information about each patient's date of transplantation and program status was entered into the system; and the nurses, pharmacists, and support staff started using the messaging application. The office staff supported this implementation by sending all phone messages electronically instead of leaving paper messages. On May 28, 2004, the alerts were implemented. By fall 2004, the three liver transplant nurses and their assistant were using the alerts to review laboratory results and identify patients overdue for testing. The medical assistant transcribed laboratory results to the paper flowchart (previously, the nurses performed this task).

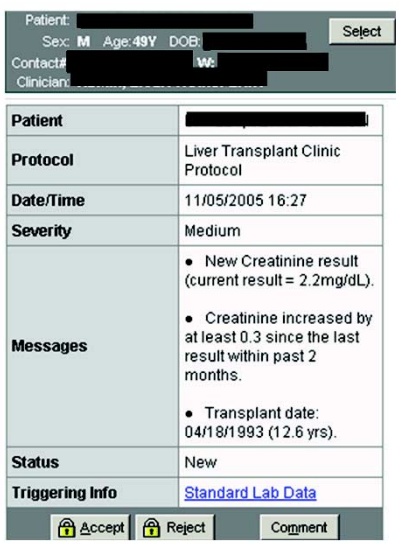


Figure 2 - Alert for a liver transplant patient

Assessment of data quality on the paper flowchart

Among results collected within 7 days prior to the chart review, the proportion of results missing on the paper flowchart decreased from 33 to 16%, before and after implementing alerts, respectively. Among results collected 14 days to 4 months prior to the chart review, the proportion of missing results decreased from 12 to 5% after implementing alerts. The change was significant for both time intervals (Pearson chi-square 0.02). The missing results may not have been transcribed to the flowchart for a variety of reasons.

Impact on workload and workflow

The computerized alerts increased workload but improved workflow. During 2005, the nurses received approximately 130 new creatinine alerts each week. In addition, each week the nurses received 12 alerts for patients newly recognized as overdue for testing. Nurses could quickly find patients with abnormal values and trends and could review laboratory results online and respond from the hospital setting where physicians were accessible for consultation.

Previously, the nurses could only manage laboratory reports while physically in the transplant office. The alerts facilitated cross-coverage among the staff. The nurses managed one list of alerts and could view all new alerts and the actions taken by others. The overdue alerts consistently identified patients that needed follow-up and led to a new system for sending letters to overdue patients. In addition, clinicians were notified when their patients became hospitalized in an IHC facility. As of December 2006, the liver transplant team continues to use the system daily to manage their patients and receive 50-60 alerts each day. They want to expand the system to manage kidney patients.

Discussion

To our knowledge, there is no published literature describing a system analysis of transplant patient care or the design of alerts for the outpatient care of liver transplant patients. The alerts were successful because we identified a problem important to the clinicians that was amenable to information technology. We fit the solution into the workflow, and created a simple intervention.[9] The clinicians have an improved system for tracking patients and responding quickly to laboratory data. In addition, the program administrator has a tool for communicating needs (Figure 1) and was able to get resources for a pre-transplant application.

The transplant patient record continues to be a hybrid of electronic and paper systems. The paper flowchart is still used because clinicians need to view laboratory results and medication dose side by side. Medications are not stored in the EHR in a manner to support this view. The alerts improved the completeness of the record, but the process of transcribing lab results creates a record that remains less than 100% complete.

Conclusion

The objectives defined for this project were met, largely because the system analysis informed us about how to integrate transplant program needs with existing tools and clinical systems. The information system has improved the quality of data available for decision- making. The requirements and systems developed for this project can be extended to meet other transplant program needs.

Acknowledgments

We thank the LDS Hospital transplant team, IHC, and the National Library of Medicine (training grant #2T15LM07124) for their support.

References

[1] UNOS. Based on organ procurement transplantation network (OPTN) data as of November 24, 2006. Accessed November 27, 2006. Available from http://www.optn.org/data/.

[2] HHS/HRSA/HSB/DOT. 2005 OPTN/SRTR Annual Report 1995-2004 (http://www.hrsa.gov/)

[3] Whitten J, Bentley L, Dittman K. System analysis and design methods. fifth ed. Boston: Irwin / McGraw-Hill 2001.

[4] Sarkar MS SD, Giuse D, Ing, Shultz E, Miller RA. Supporting longitudinal care for transplant patients with an external laboratory data entry application (poster). Proc AMIA Annu Fall Symp; 2001 Nov 3-7, 2001; Washington, DC; 2001.

[5] Gordon RD, Markus B, Mitchell S. A liver transplant center information management system. Gastroenterol Clin North Am. 1988 Mar;17(1):61-9.

[6] Kurtz M, Bennett T, Garvin P, Manuel F, Williams M, Langreder S. Demonstration of SLUMIS: a clinical database and management information system for a multi organ transplant program. Proc Annu Symp Comput Appl Med Care. 1991:889-90.

[7] Clayton PD, Narus SP, Huff SM, Pryor TA, Haug PJ, Larkin T, et al. Building a comprehensive clinical information system from components. The approach at Intermountain Health Care. Methods Inf Med. 2003;42(1):1-7.

[8] Staes CJ, Bennett ST , Evans RS, Narus SP, Huff SM, Sorensen JB. A case for manual entry of structured, coded laboratory data from multiple sources into an ambulatory electronic health record. J Am Med Inform Assoc. 2006 Jan/Feb;13(1):12-5.

[9] Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc. 2003 Nov-Dec;10(6):523-30.

Address for correspondence

Catherine Staes, BSN, MPH, PhD;
Department of Biomedical Informatics;
University of Utah
Email: Catherine.staes@hsc.utah.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

St Elsewhere's or St Everywhere's: Improving Patient Throughput in the Private Hospital Sector

Jennifer A Laffey[a] and Moran Wasson[b]

[a] *Sydney Adventist Hospital Information Services Department Consultant, EpiSoft Pty Ltd*
[b] *Sydney Adventist Hospital Director of Clinical Services, Acute Care and Perioperative*

Abstract

Communication errors have been found to be most common root cause of medical errors by the US-based Agency for Healthcare Research and Quality [1]. Although elective admissions to hospital involves a high volume of important healthcare communications where incorrect, missing or illegible information could result in a serious medical error, there is little published research on the impact of improving pre-admission communication flow between admitting doctors and hospitals.

The Sydney Adventist Hospital (the San) is a 341-bed private hospital in Sydney's northern suburbs that provides a comprehensive range of health services. A process improvement program began in early 2005 to streamline preadmission communications. The objectives of this ongoing program are broadly to improve patient safety and to increase operating efficiency. The first major initiative within this program was to implement a standardised method for inpatient booking / referral with over three hundred admitting doctors. Eighteen months on, the hospital has been able to demonstrate a significant shift in the timeliness of patient bookings from specialists' rooms, more comprehensive provision of clinical indicators that can facilitate resource planning in operating theatres and on the wards, and reduction in the ratio of bookings made in areas other than the hospital bookings department. The program continues with focus on improving accuracy of data entry, rationalising patient forms, making more effective use of information received and automation of pre-admission information flows.

Keywords:

patient admission, operating rooms, perioperative care, patient throughput, preadmission communication, inpatient booking

Introduction

The case for improving efficiency

The Australian private hospital sector comprises a substantial proportion of the Australian acute care sector with nearly 40% of all inpatient admissions and 56% of surgical admissions being to private hospitals [2]. The number of separations in the private hospital sector has grown by nearly 4% in the period 2003-4 to 2004-5 compared with a 2.6% increase in public hospital separations in the same period. There has been a 6% increase in the number of same-day separations from private hospitals in 2004-5 compared with the previous year [2].

Private hospitals have a need to more effectively manage patient throughput to meet this growing demand while optimising patient safety and satisfaction with the services provided.

The case for improving safety

Communication errors have been found to be the most common root cause of medical errors by the US-based Agency for Healthcare Research and Quality [1].

A study of 197 anaesthetic-related incidents by Kluger et al [3] found that poor communication was a major contributing factor to patient death, comorbidity or unplanned admission to ICU or HDU following surgery. Similarly, inadequate pre-admission communication was found to be a contributing factor in the murder of a woman by a psychiatric patient who was incorrectly diagnosed and discharged from a facility in the Hunter region of NSW [4].

Recent examples of poor pre-admission communication in the Sydney Adventist Hospital (the San) have resulted in:

- a medical secretary providing the wrong side information on a surgical list. This was subsequently entered to the theatre management system as received i.e. incorrectly. No written documentation was received from the doctor. This was picked up by the surgeon prior to the operation;
- no notification of a patient's recent admission to another hospital that would have alerted the San to the patient being at greater risk of having a multi-resistant organism (which they were later found to have). The patient was subsequently admitted to a shared room on the ward placing another patient at increased risk of infection

There were over 4 million elective inpatient admissions in Australia in 2004-5 (AIHW 2004-5). Elective admissions to hospital involve a high volume of important healthcare communications where incorrect, missing or illegible information could involve a sentinel event (such as death from anaesthetic complications or operation on the wrong body part).

The purpose of this research was to compare the quality of preadmission information before and after implementation of the new booking process. While a number of studies have reported on quality improvement projects in the preadmission area [5,6,7], there is little published research about the impact of improving preadmission communication flows between admitting doctors and hospitals.

About this facility

The San is a 341-bed not for profit private hospital in Sydney's northern suburbs that provides a very comprehensive range of health services including medical and surgical acute care, diagnostic services and emergency care. It runs twelveoperating theatres, two cardiac catheter laboratories, two endoscopy rooms and inpatient radiology procedure rooms. It admits about 43,000 patients per annum inclusive of 27,000 surgical, endoscopic and cardiac cathteterisation cases.

There are doctors on staff at the hospital but most of the medical care is provided by Visiting Medical Officers. Prior to this process improvement program being implemented, a wide range of preadmission referral documentation was provided by VMOs ranging from none at all (phone-call from secretary) to a comprehensive letter documenting comorbidities and current medications as well as planned treatment details. This was provided via the patients. Most typically, for surgical patients, a surgical list was written or typed by the doctor's secretary that included patient demographics, the planned procedure(s) and occasionally other details. This list was faxed to the hospital 2-3 days prior to the admission date of those patients. A written referral from the admitting doctor was also received with the patient paperwork in many cases.

About this study

A process improvement program was commenced in late 2004 to enhance the flow of information prior to hospital admission (Patient Throughput program). The objectives of this ongoing program are:

- to improve patient safety;
- have better and more timely information available for theatre and ward resource planning
- centralise bookings processes hence achieve efficiency gains
- streamline patient throughput on the day of admission and
- improve patients' satisfaction in administrative aspects of their hospital episode

The first major initiative within this program was to implement a standardised method for inpatient booking / referral with over three hundred admitting doctors (the "Hospital Booking Letter" project).

This paper describes the results of this initiative. It also describes some of the preliminary findings of other patient throughput initiatives.

Materials and methods

Agreement on standardised content for the Hospital Booking Letter involved extensive consultation with hospital staff, both clinical and administrative, as well as the Medical Advisory Committee.

The booking letter includes the patient's basic demographic information, provisional diagnosis, comorbidities, allergies and infections (with the key clinical indicators affecting order of surgical list specified), admission details (date, whether transfer from other facility), surgical procedure details if applicable, special equipment or resources required, pre-operative consults, surgical implants and diagnostic orders (pathology, radiology, ECG). The back page of the document includes a chart for medication orders on admission. Some of the Hospital Booking Letter is structured (tick-boxes) to reduce the time to fill by the doctor, to draw attention to clinical information of particular importance and to facilitate data entry.

The new process involved the admitting doctor completing this referral during the patient consultation and the doctor's secretary faxing it to a designated Toll Free number as soon as the decision was made to admit the patient and the date of admission was known. This involved a significant change of process in the doctors' rooms, not only by the doctors who had formerly provided no written documentation but also by their secretaries who were used to faxing one list of patients through shortly before the admission date. It also involved a major change of process in the hospital because there was a significantly larger volume of faxes being received. Faxes to the Toll Free number were printed in two different locations, one in the Hospital Bookings department and one in another administrative area to be sorted for use by the Pre-Admission Clinic staff to further process pre-operative consultations and diagnostic test requests.

Implementation with the admitting doctors involved site visits to a large number (over 100) admitting specialists to explain the benefits of the new process to themselves and their administrative staff, to provide them with information packs and to walk them through the key information that was useful to the hospital prior to admission. Where possible, both specialist and secretary were met with and where not possible, just the secretary.

For doctors admitting comparatively fewer patients per annum, the process was implemented via mail-out and explanatory telephone call to the secretaries and / or doctors.

In scope were all medical and surgical patients with the exception of Day Chemotherapy (who have additional information needs), Maternity patients (who attend a booking appointment with a midwife in first or second trimester) repeat visit renal dialysis patients (who are admitted by staff in the unit) and patients admitted via the Emergency Department. Gastroenterologists (Endoscopy bookings) were excluded initially but are currently being implemented on the new process.

Results

More timely information

Figure 1 below shows the notification profile (days between booking and admission) in the twelve months since February 2005 when only some doctors were implemented on the new process compared with February 2006 when most were using the new process. The notification profile of bookings in August 2006 reflects additional specialties (cardiologists) implemented on the new process since February 2006.

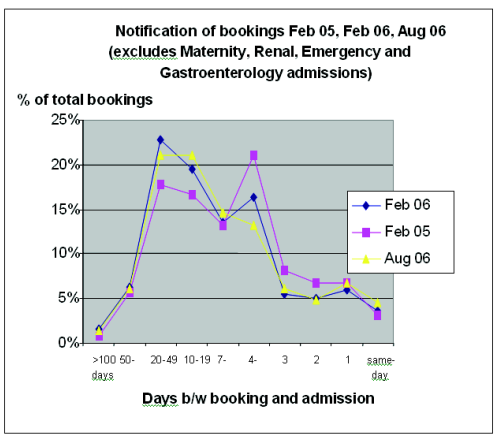


Figure 1 - Days between booking and admission

While there have been no changes in the 'tails' of the curve with booking notification greater than 50 days or on the day prior or same day of admission (or in the 7-9 day notice range), there has been a general shift towards more advance notification in the other ranges with fewer notifications being received 2-6 days prior to admission and more notifications being received 10-50 days prior to admission.

Better quality information

A random sample audit of 80 bookings (representing approximately 3% of all bookings for the period) was undertaken in the six week period from mid Sept 2006 to end Nov 2006. Bookings received via the Hospital Booking Letter contained provisional diagnosis in 86% of cases, comorbidities in 55% of cases (diabetes in 6%), allergy information (including 'nil known') in 41% of cases and infection information (including 'none') in 3% of cases. Pre-Admission Clinic (tests and / or consults) were ordered in 40% of cases.

In contrast, a sample audit of 79 bookings received in the same period via the typed surgical list method included provisional diagnosis in no cases, comorbidities in 4% of cases (diabetes in 1%), allergy information in 1% of cases and infection information in no cases.

In a smaller sample audit of 22 bookings in the same six week period, data entry accuracy was also measured. Accuracy was found to be exceptionally high for the basic surgical booking data of patient demographics, planned procedure, admission date, Medicare item numbers and surgical implants. Of the data entered, information was entered accurately albeit with a few minor spelling errors. In contrast, an unacceptably high number of errors of omission of additional data was found: provisional diagnosis and special notes/instructions were not entered at all; comorbidities were only entered in 2/6 bookings containing comorbidities (33% accuracy only) and allergies entered in 4/7 cases (57% accuracy).

Since senior staff members of the bookings department cross-check data entered on the day prior to admission, this may not have been reflective of the final level of data entry quality. However, the hospital's stated direction is that data are entered accurately and comprehensively the first time hence the audit was carried out on the initial booking entry.

Doctors' handwriting was found to be an issue in 7/22 cases (32%) in the same smaller audit sample which may have accounted for some missing data (not legible to data entry clerk).

Improved patient safety

A definitive measurement plan for this objective of the program is still being put in place making use of data from the hospital's risk management system. The key confounder to overcome is an increased number of risk incidents related to pre-admission communication arising from improved compliance in incident recording.

Centralising booking processes

The Hospital Booking Letter initiative resulted in a 6% increase in the proportion of bookings flowing to the hospital bookings department as opposed to other departments (an increase from 74% in Feb 2005 to 80% in Aug 2006).

Other objectives of the program

The Hospital Booking Letter initiative was not expected to have any effect on streamlining processes on the day of admission but nevertheless some minor benefits were noted in this area. The Pre-Admission Clinic uses the Hospital Booking Letter to identify patients who may require consultation with a case manager prior to admission. This is in addition to those patients where pre-operative tests or consults have been specifically ordered by the doctor. Attendance at the Pre-Admission Clinic reduces the time of both clerical and nursing admission on the day of admission.

With respect to improving administrative aspects of the patient's visit to the San, the more timely receipt of bookings has allowed the hospital to post out, for a larger proportion of admissions, an estimate of out of pocket expenses as opposed to providing this information on the day of admission. However, the main gains in this area are likely to arise from future patient throughput initiatives, especially the project to rationalise patient forms, implement an online pre-admission process and the project to refine the algorithm used for theatre time prediction.

Discussion

The Hospital Booking Letter project achieved a number of the stated objectives of the Patient Throughput program, particularly with respect to more timely and comprehensive receipt of pre-admission information from the admitting doctor. The main issues that arose during implementation are discussed below.

Process change

While the majority of doctors and secretaries took the process change in their stride, there were a number who were resistant to a change that they perceived as more work, both from a documentation point of view and from a process (fax per patient booking) point of view.

The project steering committee regularly reviewed a list of doctors who were not conforming, discussed the likely barriers and agreed methods to overcome these. In some cases, there were practical reasons for non-compliance (the doctor's rooms simply ran out of Booking Letter forms). In some cases, doctors with several staff in different locations did not effectively pass on details about the new process to their colleagues. In a small number of cases, there was complete refusal to conform by either the admitting doctor or their secretary or both. In general, most admitting doctors and their staff who originally did not implement the process have responded well to the rationale of improving patient safety.

Similarly the staff in the hospital took some while to get used to effectively handling and managing the additional paperwork. Reliability of the fax server was found to be of paramount importance. Given the volume of bookings received daily, a few hours' downtime of the fax server on several occasions in late 2005-early 2006 resulted in chaos in the hospital bookings department when the system came back online. The fax server was subsequently upgraded and no issues have arisen now for more than eight months.

The process is now operating efficiently with the Hospital Bookings department sorting bookings into priority order based on date of admission and entering bookings into the system within 24 hours of receipt. Similarly the administrative area located adjacent to the Pre-Admission Clinic is effectively sorting paperwork for follow-up by the Pre-Admission Clinic.

Data quality

The theatre management system routinely now contains information well in advance of patient admission about the patient's comorbidities, allergies and infections, all of which can potentially affect the order of surgical list. The Theatre Manager has been able to more effectively manage potential equipment conflicts (such as navigator probes). The Radiology Department now has increased notice of where Image Intensifiers have been ordered to plan equipment and radiographer time in theatre and ICU staff can see where an ICU bed has been ordered post-operatively. The hospital's manual handling co-ordinator has visibility of some overweight patient admissions for resource planning on the wards but it is believed that doctors under-report this comorbidity and an education program about weight information is currently underway with VMOs.

Following on from the data entry audit of late 2006, an ongoing process of monitoring and feedback to bookings staff has been put in place and a follow-up audit is planned for early 2007 to measure improvement. This process is designed to be non-threatening to individuals but to incrementally improve the accuracy of the whole department over time by focusing on problem areas in the data. The system audit trail allows for individual follow-up if required by the department manager.

Making effective use of data received

While some of the senior Nursing Unit Managers such as in day surgery and in theatres are making very good use of the information received to plan resources, others are not and require ongoing support and education about the IT system and how to more effectively use the information within it. This is partly an issue with the IT system not making key data sufficiently visible for busy clinicians to view easily and partly an issue related to nursing work practice. Both issues are currently being addressed with IT system enhancements and nurse education and training.

Current and future patient throughput initiatives

Notwithstanding some of the issues described above, the patient throughput program at the San has already achieved a number of early gains against its stated objectives with the Hospital Booking Letter initiative. Ongoing process improvement continues to more fully realise the objectives.

Patient admission forms rationalisation

During site visits to doctors' rooms for the Hospital Booking Letter project, a number of medical secretaries provided feedback that the San's patient forms were very confusing for patients. This patient paperwork pack is currently being revised into a single booklet inclusive of only essential information and forms combined. At the same time, nursing staff are working on revising the patient history form to include additional questions that would normally be asked on admission with a view to reducing the time for the nursing admission.

Pre-Admission Clinic process improvement

This service to patients continues to grow relative to overall admissions. Until recently, the scheduling process in the Pre-Admission Clinic involved a number of inefficiencies for clerical staff, especially with respect to communicating with other departments involved in the patient's visit. A scheduling system designed in house was put in place in Nov 2006. The project brief was to eliminate duplicate processes (paper and electronic diary), facilitate departmental communication and save clerical staff time. All these objectives have been demonstrably achieved within three weeks of going live with the new scheduling system.

Improving the predictive model for operating theatre times

Predicting time in the operating theatre as accurately as possible is essential for more effectively managing over- and underutilised theatres (hence streamlining patient throughput) and also improving patient satisfaction by minimising waiting time in the hospital prior to their operation.

The San currently has a model for predicting theatre times based on the procedure item number (or item number combination) and the surgeon. Basic benchmarking against other technologies in the sector has found that the San's algorithms are more sophisticated than many but they still result in a systematic underestimation of actual time in both theatre and Endoscopy procedure rooms.

It is hypothesised that other patient variables routinely captured by the hospital bookings system could also affect theatre times. The San has recently engaged a statistician to review these data and advise on whether more accurate predictions can be achieved.

Scanning received paperwork

This is being investigated as a means of improving workflow and interdepartmental communications and ensuring that essential patient paperwork (such as the consent to treat, signed by the patient and their admitting doctor) does not go missing.

Electronic automation of pre-admission

This involves a tripartite communication flow between doctor, hospital and patient and the hospital is exploring methods to facilitate online electronic booking by the doctor and online entry of admission and history information by the patient.

Conclusion

The patient throughput program at the San is a program of continuous improvement made up of a number of initiatives involving people, process and technology.

Incremental improvements in pre-admission communications have been found to have a positive impact on operating efficiencies and inpatient resource planning. Improving efficiency in the pre-admission process is essential to sustain quality of service delivery in the face of increasing demand for inpatient services.

Although improved data quality is expected to improve patient safety, the effective use of available information by clinicians is also a factor. Further research is needed to determine whether such communication improvement initiatives have a positive impact on patient safety.

Acknowledgments

The authors gratefully acknowledge Sydney Adventist Hospital for permission to publish these findings and the work of the cross-functional team involved in the Patient Throughput program especially staff in Information Services, Patient Services, Business Development, Nursing Administration, Medical Administration and Medical Records.

References

[1] AHRQ's Patient Safety Initiative: Building Foundations, Reducing Risk. Interim Report to the Senate Committee on Appropriations. AHRQ Publication No. 04-RG005. 2003 (Dec). Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/qual/pscongrpt/

[2] Australian Insitute of Health & Welfare. Australian Hospital Statistics 2004-5. Health Services Series No 26, 2006. http://www.aihw.gov.au/publications/index.cfm/title/10305

[3] Kluger MT, Tham EJ, Coleman, NA, Runciman WB, Bullock MFM. Inadequate pre-operative evaluation and preparation: a review of 197 reports from the Australian Incident Monitoring Study. Anaesthesia. 2000; 55(12): 1173.

[4] Peterson K. Where is the line to be drawn? Medical Negligence and Insanity in Hunter Area Health Service v Presland, Sydney Law Review. 2006; 28(1): 181. http://www.austlii.edu.au/au/journals/SydLRev/2006/9.html

[5] Shearer W, Monagle J, Michaels M. A model of community based, preadmission management for elective surgical patients. Can J Anaesth.1997; 44(12): 1311-1314.

[6] Hoeksema J and Munski J. Development of a Patient-Centre Preprocedure Program. AORN Journal. 1997; 65(2): 388-395.

[7] MacLeod AM. Review of a hospital preadmission program: applying the principles of continuous quality improvement. Leadership. 1997; 4(5): 11-14.

Address for correspondence

Jennifer Laffey
c/- Information Services Department,
Sydney Adventist Hospital,
185 Fox Valley Road,
Wahoonga NSW 2076
Email: jennifer.laffey@sah.org.au.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

A Meta Schema for Evidence Information in Clinical Practice Guidelines as a Basis for Decision-Making

Katharina Kaiser[a], Patrick Martini[a], Silvia Miksch[a,b], Alime Öztürk[a]

[a] *Institute of Software Technology and Interactive Systems, Vienna University of Technology, Austria*
[b] *Department of Information and Knowledge Engineering, Danube University Krems, Austria*

Abstract

Clinical practice guidelines are an important instrument to aid physicians during medical diagnosis and treatment. Currently, different guideline developing organizations try to define and integrate evidence information into such guidelines. However, the coding schemas and taxonomies used for the evidence information differ widely, which makes the use cumbersome and demanding. We explored these various schemas and developed a meta schema for the evidence information, which covers the most important components of the existing ones, is comprehensible, and easy to understand for the users. We developed and assessed the usefulness and applicability of our meta schema with guideline developers and physicians.

Keywords:

clinical practice guidelines, evidence-based medicine, recommendations, study characteristics, clinical decision support, otolaryngology

Introduction

Evidence-Based Medicine (EBM) is defined as "the integration of best research evidence with clinical expertise and patient value" [1]. EBM advocates the use of up-to-date best scientific evidence from health care research as the basis for making medical decisions. One means to communicate research evidence is to integrate it into clinical practice guidelines (CPGs), which are "systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances" [2]. Evidence-based CPGs involve a comprehensive search of the literature, an evaluation of the quality of individual studies, and recommendations that are graded to reflect the quality of the supporting evidence.

Evidence-based recommendations are mostly classified in particular grading schemas to provide a unique format at least for guidelines of the developing organization. Various organizations popularized taxonomy systems for grading the quality of evidence and the strength of recommendations (SoR) (e.g., [3-5]) and developed methodologies to categorize recommendations according to their systems.

In this paper we discuss the importance of a meta schema for levels of evidence (LoEs) and SoRs as a means for comparing, handling, and connecting LoEs and SoRs of different taxonomy systems for supporting the medical decision-making process. Due to the profusion of grading schemas users are often puzzled by the message a grade conveys. The different application of codes (e.g., I, II, III, …; A, B, C, …; 1, 2, 3, …; Ia, Ib, IIa, …) and the different definitions of the levels are not only confusing to users, but also aggravates a comparison and decreases the transparency of the schemas [6]. Table 1 and Table 2 give two examples of LoEs and SoRs from two different guideline-developing organizations.

Table 1 - Levels of Evidence used by the University of Michigan Health System

Level	Definition
A	Randomized controlled trials
B	Controlled trials, no randomization
C	Observational trials
D	Opinion or expert panel

The overall objective of this work is to facilitate the decision-making process on the basis of a systematic representation of the evidence information. A systematic representation is required to handle evidence information in computer-interpretable guideline representation languages (see [7] and [8] for a comprehensible overview). To achieve this objective we meet the following more specific objectives:

1. Development of a meta schema for grading evidence information. This schema should cover the most important components of various rating schemas for LoEs and SoRs.
2. Mapping of different LoEs and SoRs used by different organizations into this meta schema.

For our research we used 21 evidence-based CPGs from the clinical specialty otolaryngology. We selected the clinical specialty otolaryngology, because there are many, well structured guidelines available for our purpose. Based on the different LoEs and SoRs in these CPGs we have developed a meta schema for both graded and ungraded evidence information and SoRs.

Table 2 - Strength of Recommendations defined by the Scottish Intercollegiate Guidelines Network (SIGN)

Strength	Definition
A	At least one meta analysis, systematic review, or RCT* rated as 1++, and directly applicable to the target population; or A systematic review of RCTs or a body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall consistency of results
B	A body of evidence including studies rated as 2++, directly applicable to the target population, and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 1++ or 1+
C	A body of evidence including studies rated as 2+, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2++
D	Evidence level 3 or 4; or Extrapolated evidence from studies rated as 2+

* RCT … Randomized Controlled Trial

The meta schema connects existing grading systems to provide a means to increase the transparency among the various schemas and to appraise ungraded recommendations. By the direct comparability of various grading systems the communication of the underlying information is quickly and concisely possible.

Furthermore, decision-support is a crucial topic in computer-supported guideline's research. The meta schema will thereby facilitate the integration of evidence information and form a basis to handle the multitude of grading systems on an equal level.

In the following section we describe the process of developing our meta schema and we show the results in the subsequent section. Furthermore, we discuss the outcomes and cease with concluding remarks and future work.

Methods

Guyatt et al. [9] defined several criteria for developing an optimal grading schema. As our intention is not to develop a new system, but a meta schema that connects the existing schemas, other requirements apply. These are to have sufficient categories to cover a good portion of systems, to be consistent with existing systems, and to be simple and transparent to users.

We selected 21 CPGs from the National Guideline Clearinghouse[1], which were developed by nine different organizations (see Table 3). These CPGs cover eight different representations of LoEs and three different representations of SoRs.

Table 3 - Guideline development organizations

Organization/Cooperation	Number of Guidelines
American Academy of Family Physicians; American Academy of Otolaryngology – Head and Neck Surgery; American Academy of Pediatrics	1
American Academy of Pediatrics	2
Allergic Rhinitis and its Impact on Asthma Workshop Group	1
Cincinnati Children's Hospital Medical Center	3
Finnish Medical Society Duodecim	1
Institute for Clinical Systems Improvement	6
Practice Guidelines Initiative	1
Scottish Intercollegiate Guidelines Network	3
University of Michigan Health System – Academic Institution	3

After analyzing the different grading schemas of these organizations we decided to use the definitions of LoEs and SoRs of the Scottish Intercollegiate Guidelines Network[2] (SIGN) [10] as a basis for our meta schema. One reason for this decision was that SIGN's representation of evidence information is systematically evaluated and clearly structured and defined. As the SIGN approach does not cover all required information to represent the different LoEs and SoRs, we expanded the definitions for other organizations to cover their representations of the evidence information, too.

The most relevant attributes for developing the meta schema are:

- Code schema of guideline developing organization
- Study design and quality
- Strength of recommendations
- Benefits and harms

During the development process we conducted interviews with various guideline developers and physicians. We discussed the correctness, sensibility, availability, and understandability of the hierarchical structure, the quality of the LoEs and SoRs, the mapping tables, and the balance between benefits and harms. Furthermore, we surveyed the

1 http://www.guidelines.gov (last assessed December 3, 2006)
2 http://www.sign.ac.uk (last assessed December 1, 2006)

availability of required information and the facilitation of the decision-making process. Moreover, the covering of the meta schema with existing grading systems was verified during the entire process. The remarks and comments were incorporated in our schema altogether.

Code schema of the guideline developing organization

This attribute is essential to differentiate between various grading schemas, because a symbol or code communicating a grade can represent different meanings (see also [6]). For example, the *University of Michigan Health System* uses the symbols "*A", "B", "C", "D*" for LoEs, whereas *SIGN* uses these symbols for SoRs. Thus, it is not possible to extract the evidence information from the guidelines and map them to the meta schema, without the information about the developing organization.

Study design and quality

The quality of evidence is described by LoEs. They are mostly explicitly represented in the guidelines but different symbols are used to refer to them (see for instance Table 1).

The attribute of the study design covers all study types (i.e., Meta-Analysis, Systematic Reviews, Randomized Controlled Trials (RCTs), Cohort Studies, Case Control Studies, Expert Opinion) used in CPGs.

We represent the LoE on the basis of the study design's attribute, because in that way we get an ordered structure, where *meta-analysis* is on the top of the hierarchy and *no study design* is at the bottom.

The study design attribute plays a significant role by assigning a grade to ungraded evidence information. Often CPGs include information about the study design upon which the recommendations are based, but they do not provide any explicit grades for their evidence (e.g., "*The recommendations are supported by randomized controlled trials. Adverse parasympathetic events were reported by participants in randomized controlled trials, the most frequent and troublesome being increases sweating which occurred in about one-quarter of patients taking 5 mg three times per day and about one-half of patients taking 10 mg*" [11]).

Another attribute to be considered for establishing the LoEs is the study quality. It refers to the detailed study methods and execution. The study quality is thereby the degree to which a study employs measures to minimise biases, focussing on internal validity [12].

Our representation has to address both the study design and the study quality. The levels have to be clearly distinguishable and easily and clearly interpretable.

Strength of recommendations

For SoRs also different symbols or names are used but they are not always explicitly mentioned in our CPGs. Three out of nine organizations have defined SoRs and only six of the 21 CPGs include explicitly defined SoRs. In 15 CPGs no information about SoRs is included.

The representation of SoRs has to be representable to the different existing SoRs. The requirements for our SoR taxonomy are that

- The strengths have to be clearly distinguishable from each other
- The names of the grades have to be meaningful
- The strengths have to be easily and clearly interpretable
- The number of grades should be limited to ensure an easy understanding and application

The GRADE Working Group[3] developed a system for defining the recommendations based on four factors [4]:

1. The trade-offs, taking into account the estimated size of the effect for the main outcomes, the confidence limits around those estimates, and the relative value placed on each outcome
2. The quality of the evidence
3. Translation of the evidence into practice in a specific setting, taking into consideration important factors that could be expected to modify the size of the expected effects, such as proximity to a hospital or availability of necessary expertise
4. Uncertainty about baseline risk for the population of interest.

Recently, medical associations and organizations adapt the GRADE approach for their needs [9,13]. However, the publication of guidelines using the new grading systems will take time. In our CPGs, this new approach is not implemented yet. The strength of a recommendation is only based on the underlying quality of evidence.

Benefits and harms

Information about benefits and harms of a particular treatment plays a significant role in the decision-making process. But in our guidelines they are described very briefly and limited. They do not contain information about the trade-off between the benefits and harms either. For embedding information about benefits and harms into the decision-making process, we need them to be represented explicitly. Thus, a schema for the trade-off between benefits and harms is necessary, because this information is essential for the medical staff to assess benefits and harms of a treatment recommendation. For example:

1. In patients with peptic ulcer, drug A reduces acidity. This recommendation is based on RCTs.
2. In patients with cardiac problems, drug A may cause heart attacks and hence is contraindicated. This recommendation is based on case reports.

The second argument is based on lower quality evidence, but defeats the first argument, because of the more important claim (heart attack is worse than having acidity reduced).

Results

The meta schema

Based on the considerations described in the previous section we developed a meta schema for representations of the quality of evidence, the strength of recommendations, and benefits and harms.

3 http://www.gradeworkinggroup.org (last assessed November 30, 2006)

Definition of levels of evidence

Our definition of LoEs is based on the *study design* and the *study quality*. The *study design* is essential to get a hierarchical representation and to assign a level to ungraded evidence information in CPGs. The LoEs consist of symbols that cover information about the study design and the quality of the studies. We introduced our own symbols (e.g., I_1, I_2, … II_1, II_2, …) that represent both the study design and quality. The first character describes the study design whereas the number describes the quality. Table 4 shows a part of our LoEs schema.

Table 4 - Part of the meta schema representing Levels of Evidence

Study Design	Evidence Level	Definition
Meta Analysis	I_1	Meta-analysis of RCTs
	I_2	High quality meta-analysis
	I_3	Well-conducted meta-analysis
	I_4	Meta-analysis
Systematic Reviews	II_1	High quality systematic reviews of RCTs with large sample
	II_2	High quality systematic reviews of RCTs with small sample
	II_3	High quality systematic reviews of RCTs with very low risk of bias
	II_4	Systematic reviews of RCTs
	II_5	High quality systematic reviews of cohort studies
	II_6	High quality systematic reviews of case-control studies
	II_7	Systematic reviews
…	…	…

Definition of strengths of recommendations

Based on the requirements for our SoRs taxonomy we defined four different strengths, because more than four hierarchical levels are hardly distinguishable, but less do not adequately cover existing systems. The four grades are:

1. Strong Recommendation
2. Recommendation
3. Weak Recommendation
4. No Recommendation

They have a unique definition and are easy to differentiate from each other. For example, *Strong Recommendation* is directly applicable to the target population and bases on at least one meta-analysis, systematic review of RCTs, RCTs with very low risk of bias, high quality meta-analysis of observational studies, or high quality systematic reviews of observational studies.

Our aim with these definitions of SoRs is to provide guideline users a proposed recommendation that should only be a direction if there is no explicit representation of SoRs in the CPGs.

Definition of trade-off between benefits and harms

Table 5 shows the definitions of the trade-off between the benefits and harms. Our definitions are based on the descriptions used by the GRADE working group [4], because they have a well defined categorization of the trade-off between the benefits and harms in their grading schema.

Table 5 - Schema for trade-off between benefits and harms

Classification	Benefits and Harms
Clear Benefit	The benefits of the recommended approach clearly exceed the harms.
Benefit	The recommended intervention explicitly does more good than harm or the benefits outweigh the harms.
Unclear Balance	It is unclear whether the recommended intervention does more good than harm. The trade-off between benefits and harms is quite unclear.
No Clear Benefit	The recommended intervention clearly does not do more good than harm.

Mapping the evidence information of CPGs

The meta schema should provide a general representation of different classifications of LoEs and SoRs used in CPGs. We connected each individual LoE and SoR taxonomy to our meta schema. Figure 1 shows a mapping between SIGN and our meta schema. The mapping tables provide guideline users a better handling and understanding of the evidence information in CPGs. With this representation the users have a means to compare different LoEs and SoRs.

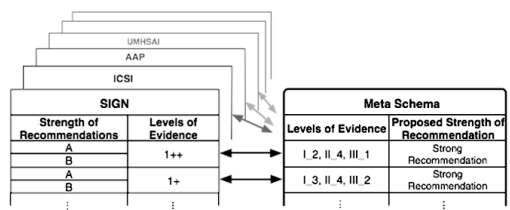


Figure 1-– Mapping table for the meta schema showing mappings from and to the SIGN schema

Discussion

The most significant factor for decision-making is the strength of recommendations. In most taxonomies the following aspects are taken into account:

- The level of evidence of individual studies
- The type of outcomes measured by these studies (patient-oriented or disease-oriented)
- The number, consistency, and coherence of the evidence as a whole
- The relationship between benefits, harms, and costs

In most guidelines this information is not entirely available. Thus, it is only possible to assign a constricted and intermediate SoR. Furthermore, benefits and harms for each recommendation are needed to incorporate them in the constitution of the SoR. But often, benefits and harms are only given for the entire guideline and not individual recommendations.

Conclusion

Our meta schema is an instrument to connect different schemas of LoEs and SoRs. The meta schema is representable to eight different systems defining LoEs and three different systems defining SoRs and incorporate the ideas and concepts of the GRADE Working Group. Furthermore, it is possible to assign a LoE to an ungraded evidence recommendation based on the study design and quality if available. It covers also information about the trade-off between benefits and harms, which are mostly not included in the existing grading schemas. We used the attributes study design and study quality (defined in LoEs), SoRs, the organization's code schema, and benefits and harms, which were significant for the development process.

Furthermore, we think that our meta schema can also support instruments for guideline appraisal (e.g., AGREE [14], GLIA) in terms of providing means to better understand and compare the various existing grading schemes for evidence information.

Acknowledgments

This work is supported by "Fonds zur Förderung der wissenschaftlichen Forschung - FWF" (Austrian Science Fund), grants P15467-N04 and L290-N04.

References

[1] Sackett D. Evidence Based Medicine: What it is and what it is not. BMJ, 312:71, 1996.

[2] Field M and Lohr K. Clinical Practice Guidelines: Directions for a New Program. National Academy Press, 1990.

[3] Harbour R and Miller J. A new system for grading recommendations in evidence based guidelines. BMJ, 323:334–336, 2001.

[4] GRADE Working Group. Grading Quality of Evidence and Strength of Recommendations. BMJ, 328(7454):1490-1498, 2004.

[5] Ebell MH, Siwek J, Weiss BD, Woole SH, Susman J, Ewigman B, and Bowman M. Strength of Recommendation Taxonomy (SORT): A Patient-Centered Approach to Grading Evidence in the Medical Literature, American Family Physician, 69(3): 548–556, 2004.

[6] Schünemann HJ, Best D, Vist GE, and Oxman AD for the GRADE Working Group. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. Canadian Medical Association Journal (CMAJ), 169(7): 677–680, 2003.

[7] Peleg M, Tu SW, Bury J, Ciccarese P, Fox J, Greenes RA, Hall R, Johnson PD, Jones N, Kumar A, Miksch S, Quaglini S, Seyfang A, Shortliff E, and Stefanelli M. Comparing Computer-Interpretable Guideline Models: A Case-Study Approach, JAMIA, 10(1):52–68, 2003.

[8] de Clercq PA, Blom JA, Korsten HHM, and Hasman A. Approaches for Creating Computer-interpretable Guidelines that Facilitate Decision Support, Artificial Intelligence in Medicine, 31(1):1–27, 2004.

[9] Guyatt GH, Guttermann D, Baumann MH, Addrizzo-Harris D, Hylek E, Philips B, Raskob G, Zelman Lewis S., and Schünemann HJ. Grading strength of recommendations and quality of evidence in clinical guidelines: report from an American college of chest physicians task force. Chest, 129:174–181, 2006.

[10] Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developers' handbook. SIGN, 2004.

[11] Practice Guideline Initiative. Symptomatic treatment of radiation induced xerostomia in head and neck cancer patients, 2004.

[12] Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, and Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. Control Clin Trials, 16:62-73, 1995.

[13] Schünemann HJ, Jaeschke R, Cook DJ, Bria WF, El-Solh AA, Ernst A, Fahy BF, Gould MK, Horan KL, Krishnan JA, Manthous CA, Maurer JR, McNicholas WT, Oxman AD, Rubenfeld G, and Turino GM. An Official ATS Statement: Grading the Quality of Evidence and Strength of Recommendations in ATS Guidelines and Recommendations. American Journal of Respiratory and Critical Care Medicine, 174:605-614, 2006.

[14] The AGREE Collaboration. Writing Group: Cluzeau FA, Burgers JS, Brouwers M, Grol R, Mäkelä M, Littlejohns P, Grimshaw J, Hunt C. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Quality and Safety in Health Care 2003; 12(1): 18–23.

[15] Shiffman RN, Dixon J, Brandt C, Essaihi A, Hsiao A, Michel G, and O'Connell, R. The guideline implementability appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. BMC Med Inform Decis Mak, 5, July 2005.

Address for correspondence

Katharina Kaiser
Vienna University of Technology
Institute of Software Technology & Interactive Systems
Favoritenstrasse 9-11/188, A-1040 Vienna, Austria
Email: kaiser@ifs.tuwien.ac.at

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Creating Interoperable Guidelines: Requirements of Vocabulary Standards in Immunization Decision Support

Karen M. Hrabak MSN RNC[a], James R. Campbell MD[b], Samson W. Tu MS[c], Robert McClure MD[d], Robert (Tony) Weida PhD[d]

[a] *The Nebraska Medical Center, United States*
[b] *University of Nebraska Medical Center, United States*
[c] *Stanford Medical Informatics, Stanford University, United States*
[d] *Apelon, Inc., United States*

Abstract

Interoperable support of electronic health records and clinical decision support technology are central to the vision of sustainable information infrastructure. Efforts to implement interoperable clinical guidelines for immunization practice have been sparse. We used the SAGE knowledge workbench to develop a knowledge base to provide immunization decision support in primary care. We translated the written clinical guideline into a structured decision logic format. The semantic content to completely capture CDC clinical decision logic required 197 separate concepts but was completely captured with SNOMED CT and LOINC. Although 88% of concepts employed pre-coordinated codes, 6% of guideline concepts required expanded vocabulary services employing Boolean logical definition using two or more SNOMED concepts. Post-coordination requirements were modest, representing just 6% of guideline semantic concepts. We conclude that creation of interoperable knowledge bases employing clinical vocabulary standards is achievable and realistic. Employment of information model (HL7 RIM) and vocabulary (SNOMED CT, LOINC) standards is a necessary and feasible requirement to achieve interoperability in clinical decision support.

Keywords:

interoperability, standard vocabulary, clinical practice guidelines, clinical decision support

Introduction

An expert system employs an inference method linked to a domain ontology [1], ideally evaluates patient state data from the electronic health record, and issues recommendations for care. Advances in clinical vocabulary development [2] and interest multi-nationally [3] has promulgated a core set of reference terminologies which offer to evolve into the comprehensive clinical ontologies needed for electronic health record (EHR) technology and decision support. SNOMED CT is a core reference terminology within these recommendations. NDF-RT and laboratory LOINC are controlled vocabularies with reference features also within the core of standards recommended within the US. At this time however, implementation of these standards is infrequent, in part related to confusion regarding best deployment and difficulties with conversion of legacy data. While a single study report [4] has argued that these terminologies may be insufficient to support guidelines, the functionality of these terminologies relative to representation of domain knowledge for a guideline expert engine poses complex issues not yet resolved.

Clinical practice guidelines seek to standardize care and facilitate the provision of evidence-based care. Historically published in free text formats, efforts to encode and implement guidelines within the EHR for clinical decision support face many challenges. Interpretation and translation of the written guideline text is necessary. Ambiguity within the source publication requires clinical expertise to precisely formulate the guideline logic [5]. Concept modeling problems exposed by guideline encoding include differences in granularity and definition between the guideline and the domain ontology and interactions with the vendor information model.

The National Vaccine Advisory Committee guides best practices surrounding immunization administration in the United States [6]. Included in their recommendations is an emphasis on accurate vaccine recording practices and a method to send clinical reminders to patient and practitioners. Other investigators have recognized the importance of this domain and have worked to create decision support for immunization practice. These past efforts have successfully modeled immunization clinical decision support forecasting and reminder systems [7-8]. The IMM/SERV system supported childhood immunization forecasting and maintained a web based knowledge maintenance and testing environment [9]. The immunization reminder recall system [8] provided immunization decision support utilizing a modular architecture. Concepts in its knowledge base have been mapped to medical entities dictionary (MED) employed at the New York Presbyterian Hospital to provide integration with clinical records at that facility. Representation and maintenance of the knowledge

domains in these systems employed tabular, rules based and procedural approach. Neither of these past implementations has employed clinical vocabulary standards. The resulting decision models employed were unique to the environments in which they were developed.

The SAGE [10] consortium is a collaboration of academic and private sector interests with the shared goal of creating interoperable guideline decision support. In order to support interoperability, the knowledge bases created for SAGE employ only the core

vocabulary resources recommended by the US National Committee on Vital and Health Statistics (NCVHS). Knowledge modeling occurs on an open source workbench created with the Protégé [11] knowledge tool.

The program interface between the SAGE inference engine and the vendor clinical information system communicates via an HL7 reference information model (RIM) compliant query engine termed the virtual medical record (vMR). For purposes of inference support the SAGE engine employs a suite of vocabulary services which bind the decision support software to the SNOMED CT structures. This binding provides ontologic features of subsumption and concept definition. Concept definition beyond the pre-coordinated scope of SNOMED CT is handled with vocabulary service extensions and post-coordination within the SAGE SNOMED CT extension. The end result is a knowledge construction and domain ontology (extension of SNOMED) which can freely interoperate with any other system compliant with SAGE and SNOMED CT standards.

Given the clinical importance of immunization practices and the historical efforts of other decision support scientists, achieving interoperability through the use of standard terminologies is critical. We therefore organized, enumerated and characterized the vocabulary and knowledge services of the SAGE immunization guideline in order to inform the concerns of EHR vendors and emphasize the benefits of vocabulary standards compliance.

Methods

Guideline clarification and logic modeling

The US Center for Disease Control (CDC) through the Advisory Committee on Immunization Practices issues guidelines for the US with specific recommendations for vaccination of child and adult populations [12]. We obtained the immunization recommendations (Fall 2005) for children and adults for this analysis. Our initial task was to translate by hand the written guideline content into a structured representation. We compiled indications, contraindications and deferral criteria for each vaccine. Age specific criteria for eligibility, dosing intervals, and catch up rules for missed vaccine doses were identified. These criteria were used to formulate a knowledge base specification document containing logical IF-THEN statements which formalized the CDC logic while employing the source guideline concept statements. Figure 1 represents an example of a logic statement derived from the pediatric immunization guideline.

From the logic base we compiled an inventory of concept references employing methods we have described [13] and worked with local clinical experts to disambiguate guideline statements which were unclear. For example, the guideline source concept ‘progressive neurologic disorder’ required clinical domain expert definition, resulting in a logical union of the concepts ‘Lennox Gastaut’, ‘Tuberous sclerosis’ and the nested union of concepts ‘Developmental delay’ and ‘Encephalopathy’.

We then compared our concept inventory against pre-coordinated concepts from SNOMED CT and LOINC. When source concepts were not pre-coordinated but could be formulated correctly with logical constructions of pre-coordinated concepts, we did so. Remaining concepts that were clearly outside of the scope of pre-coordinated SNOMED

CT were modeled into an extension namespace for the guideline following editorial principles published by the College of American Pathologists [14].

Rule 1: First dose MMR pediatric	
IF	NO CONTRAINDICATION TO MMR
	AND
	NO REASON FOR DEFERRAL
	AND
	AGE >= 12 MONTHS
	AND
	NUMBER OF MMR DOSES = 0
	AND
	NO VARICELLA VACCINE ADMINISTERED WITHIN 28 DAYS
THEN	
	ADVISE ADMINISTER MMR VACCINE

Figure 1 – Structured decision logic statement

All immunization logic rules and vocabulary concepts were then linked to a set of EHR queries employing the idealized record structure (vMR) which we have developed with the HL7 Clinical Decision Support Technical Committee in compliance with the HL7 Reference Information Model (RIM) [10]. This link involved an explicit assertion of the expected EHR records required and the

attributes required to bind the decision logic to the clinical patient record. Detailed data models of guideline concepts identify how patient data associated with the concepts are represented by a vMR class. Table 2 lists the 13 primary classes of the vMR in the left hand column.

Vocabulary services

Recognizing that guideline statements sometimes requested unique concepts for query from the record, while other statements implied retrieval from within a set of concepts (all instances of diabetes), we reviewed the guideline to clarify the vocabulary services required. We categorized all vocabulary service requirements on a scale reflecting the complexity of the concept relative to pre-coordinated NCVHS vocabularies, and the expected query function in the run-time environment.

Categorization of concept query requirements

Category 1: Concept instance only is directly referenced by the guideline logic. This category includes instances such as gender code, qualifier values (e.g. contraindicated or true) and lab codes (e.g. Hepatitis B surface antigen, Measles virus IGG antibody).

Category 2: Concept, along with all specializations, are implicitly referenced by the guideline. This category includes concepts that may have many conceptual variations within the EHR (such as "Diabetes mellitus" or "Hemoglobinopathy") and the guideline expects all more specialized children of the concept to be included at run-time query.

Category 3: Boolean constructions

- 3a: Guideline concept is represented by the logical Boolean 'OR' (Set union) of two or more category 2 references. For example "Functional or anatomic asplenia" is logically defined by the union of the sets of concepts: "Splenectomy", "Functional splenectomy", "Congenital asplenia", "Sickle cell disease", "Asplenia syndrome" and "Hyposplenism" - including children - within the SNOMED CT ontology.
- 3b: Concept requires the logical Boolean 'AND' (Set intersection). An example of a category 3b concept is "bisexual male" which is the intersection of the concept sets "Patient is male" and "Bisexual".
- 3c: Concept expression includes a Boolean 'NOT' (Set complement). Concepts in this category include concept expressions that exclude descendant concepts within a hierarchy. For example, the guideline concept of "Chronic respiratory disease" when clinically reviewed, was defined to exclude the SNOMED specialization concepts of "Chronic rhinitis" and "Chronic sinusitis" at run-time.

Category 4: Concept post-coordination required. This category includes concepts requiring extension development for SNOMED CT. Since they are not pre-coordinated and cannot be defined from logical statements employing pre-coordinated concepts, they are defined employing SNOMED formalisms as extension vocabulary. An example for the concept "Maternal hepatitis B surface antigen positive" is included in Figure 2.

SAGE knowledge modeling

Employing the immunization rule logic and the concept inventory, we then proceeded to model the complete guideline using the SAGE guideline ontology [10]. Context of care, clinical workflow and organizational resources are elements of the SAGE ontology. All decision logic rules and vocabulary queries were bound and modeled employing SAGE formal criteria and SAGE actions linked to the vendor EHR. For comparison with previous immunization guideline work, we counted and summarized these execution elements.

The SAGE guideline workbench produces an XML knowledge base that can be shared between clinical systems. We validated the immunization knowledge base with a series of experiments including syntax checking of the XML and simulated run-time assessment employing test cases. We are now validating the knowledge base against actual clinical records in two separate enterprise systems.

Results

Vocabulary inventory

This "birth-to-death" immunization knowledge module was a complex construction. The 45 pages of clinical guideline publication were distilled into 75 separate "IF-THEN" statements in support of three clinical implementation scenarios proposed by the clinical team. The scenarios included vaccination advice at birth, a primary care office visit, and a survey scenario for population based reminders. An inventory of the source utterances from the guideline statements yielded 147 conceptual references. Disambiguation and expert clinical opinion was required with 7 concepts which were then defined within the SNOMED CT extension ontology. Table 1 provides a summarization of the pre-coordinated vocabulary concepts that were ultimately required to support guideline logic. These concepts were installed in vMR queries as data type restrictions which defined the value sets for retrieval of information from the EHR by the SAGE decision engine. Each query employed one or several coded concepts from the distinct concepts tallied for the guideline.

```
<concept conceptID= "12110001000004100"
  conceptStatus="0"  fullySpecifiedName="Maternal
   hepatitis B surface antigen  positive (situation)"
   isPrimitive="0">
  <descriptions>
    <description term="Maternal hepatitis B surface
   antigen
     positive(situation)"   descriptionStatus="0"
     descriptionType="3"  languageCode="en-US"
     initialCapitalStatus="0"
  </descriptions>
  <relationshipSet>
    <relationship relationshipType="116680003 IS-A"
       conceptId2="4164710007 Family history of
       clinical finding"  characteristicType="0"
       refinability="0"/>
    <relationshipGroup>
```

```
    <relationship relationshipType="246090004
   Associated
       finding" conceptId2="165806002 Hepatitis B
       surface
       antigen positive" characteristicType="0"
        refinability="0"/>
    <relationship relationshipType="408729009 Finding
      context" conceptId2="410515003 Known present"
       characteristicType="0" refinability="0"/>
    <relationship relationshipType"408731000 Temporal
       context" conceptId2="410512000 Current or
       specified"
       characteristicType="0" refinability="0"/>
     <relationship relationshipType"408732007 Subject
        relationship context" conceptId2="72705000
       Mother" characteristicType="0" refinability"0"/>
     </relationshipGroup>
  </relationshipSet>
</concept>
```

Figure 2 – Post-coordinated concept definition

Table 2 summarizes the final analysis of vMR queries required to support rule logic with a total of the vocabulary concepts. Not all guideline concepts could be accurately modeled employing pre-coordinated SNOMED CT. Table 3 summarizes the complexity of the vocabulary model and services required to support immunization guidelines. This reflects run-time management requirements (only category 1 concept references do not require retrieval of data sets which include all children of the concept) as well as the requirements for post-coordinated vocabulary development (category 4 concepts).

Table 1- Pre-coordinated concepts by semantic type (Category 1 and 2 concept complexity)

SNOMED domain	***n***
Context-dependent category	3
Disorder	51
Finding	23
Observable entity	10
Occupation	4
Organism	1
Person	1
Procedure	9
Product (clinical drug)	50
Qualifier	16
Racial group	1
Substance	5
Total	*n* = 174

Table 2- Concept inventory by vMR query class

vMR query	**SNOMED**	**LOINC**
Adverse reaction	42	
Agent	0	
Alert	0	
Appointment	0	
Encounter	0	
Goal	0	
Observation	39	5
Order	6	
Medication Order	24	
Problem	78	
Procedure	4	
Referral	0	
Substance administration	32	

Characteristics of knowledge model

SAGE employs the frame-based knowledge modeling of the Protégé environment. Immunization logic criteria are formulated into frames which enforce a set of constraints on data query from the EHR. Criteria are employed within decision models which reproduce the source guideline logic and communicate with the vendor record via action specifications. The full immunization knowledge model required 236 Boolean criteria, 207 presence criteria and 161 comparison criteria. These were employed in 88 decision models which employed 82 action specifications. The full immunization guideline model, including SAGE workbench and vocabulary coding, is available for public use and evaluation [15].

Table 3- Concept inventory by complexity

Category	***n***
Category 1 (Concept entity)	35 (17.8%)

Category	***n***
Category 2 (Subsumption)	139 (70.5%)
Category 3a (Boolean with OR)	8 (4.1%)
Category 3b (Boolean with AND)	2 (1.0%)
Category 3c (Boolean with negation)	2 (1.0%)
Category 4 (Post coordination)	11 (5.6%)
Total	$n = 197$

Discussion

The development of reference terminologies for clinical vocabulary standards has created utility but also poses new challenges for the knowledge information specialist. Previous studies [16] have documented the limitations of pre-coordinated terminologies, but a commitment to compositional forms means that procedures and methods for management of post-coordination must be developed.

In contrast to a previous report [4], we found that comprehensive coding in support of our guideline was feasible, but that vocabulary services for the guideline engine had to be extended to include support for two services: 1) Boolean definitions of complex concepts and 2) integration of post-coordination within a SNOMED extension. Since management of large extension vocabulary sets requires new skills and software functionality such as description logic classifiers, this is a matter of developing understanding within the informatics community.

Run-time support provided by reference terminologies such as SNOMED CT is also important to decision support engines. Our experience clearly documents the significance of support for aggregation within record query activity. 80-90% of queries into the EHR were searching not just for a single concept, but for one within a related set. By providing for identification of all specializations of a concept with hierarchical relationships, SNOMED CT supplies knowledge structures which replace the need for exhaustive code list generation in knowledge bases. This defines a clear benefit resulting from standard reference terminology deployment, as well as an important use case for evaluating evolution of these vocabulary systems.

Conclusions

It is feasible to implement guideline decision support within a knowledge engine employing international standard vocabularies. Effective use of these reference terminologies requires new procedures for vocabulary management and deployment. Benefits to the knowledge engineer include savings in domain knowledge development and true semantic interoperability.Acknowledgement

The SAGE project has been supported by grant 70NANB1H3049 of the NIST Advanced Technology Program.

References

[1] Musen MA. Scalable architectures for clinical decision support. *Methods Inf Med.* 1999;38:229-38.
[2] Wang AY, Sable JH, Spackman K. The SNOMED clinical terms development process: refinement and analysis of content. *Proc AMIA Symp*. 2002;: 845-9.
[3] SNOMED Standards Development Organization [Online]. 2006 Nov 28 [cited 2006 Dec 1]; Available from: URL: http://www,snomed.org/news/documents/snomed_meeting_statement_nov2006_update.pdf
[4] Dykes PC, Currie LM, Cimino JJ. Adequacy of evolving national standardized terminologies for interdisciplinary coded concepts in an automated clinical pathway. *J Biomed Inform* 2003. 36:313-25.
[5] Tierney WM. Improving clinical decisions and outcomes with information: a review *Int J Med Inform.* 2001 Jun;62(1):1-9.
[6] Poland G, Shefer A, McCauley M, Webster P, Whitley-Williams P, Peter G. Standards for adult immunization practices. *Am J Prev Med.* 2003;25(2):144-150.
[7] Miller PL, Frawley SJ, Sayward FG. Issues in computer-based decision support in public health illustrated using projects involving childhood immunization. *J Public Health Manag Pract.* 2001;7(6), 75-86.
[8] Wang D, Jenders R, Dasgupta B. Translating national childhood immunization guidelines to a computer-based reminder based reminder recall system within an immunization registry. *Proc AMIA Symp*. 1999;:994-8.
[9] Miller PL, Frawley SJ, Brandt C, Sayward FG. A prototype website for immunization maintenance website. *Proc AMIA Symp*.1997;:293-7.
[10] Tu SW, Campbell J, Musen MA. The SAGE guideline modeling: motivation and methodology. *Stud Health Technol Inform.* 2004;101:167-71.
[11] Gennari JH, Musen MA, Fergerson RW, Grosso WE, Crubezy M, Eriksson H, Noy NF, Tu SW. The evolution of protégé: An environment for knowledge-based systems development. *Int J Hum Comput Stud.* 2003;58(1):89-123.
[12] Center for Disease Control, Advisory Committee on Immunization Practice [Online]. 2006 [cited 2006 Nov 28]; Available from: URL: http://www.cdc.gov/nip/ACIP/default.htm
[13] Tu S, Musen MA, Shankar R, Campbell J, Hrabak K, McClay J, Huff S, McClure R, Parker C, Rocha R, Abarbanel R, Beard N, Glasgow J, Mansfield G, Ram P, Ye Q, Mays E, Weida T, Chute C, McDonald K, Mohr D, Nyman M, Scheitel S, Solbrig H, Zill D, Goldstein M. Modeling guidelines for integration into clinical workflow. *Medinfo.* 2004; 11(Part 1):174-8.
[14] SNOMED Clinical terms Concept modeling/Style authoring guide. College of American Pathologists. October, 2004.
[15] SAGE Project Overview [Online]. 2006 Dec 3 [cited 2006 Dec 3]; Available from: URL: http://sage.wherever.org/projectoverview/projectoverview.html
[16] Campbell JR, Carpenter P, Sneiderman C, Cohn S, Chute CG, Warren J. Phase II evaluation of clinical coding schemes: completeness, taxonomy, mapping, definitions, and clarity. *J Am Med Inform Assoc*. 1997 May-Jun;4(3):238-51.

Address for Correspondence:

Karen Hrabak MSN, RNC
khrabak@nebraskamed.com

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Automatic Treatment of Temporal Issues in Clinical Guidelines in the GLARE System

Luca Anselma[a], Paolo Terenziani[b], Stefania Montani[b], Alessio Bottrighi[b]

[a] *Dipartimento di Informatica, Università di Torino, Torino, Italy*
[b] *Dipartimento di Informatica, Università del Piemonte Orientale "Amedeo Avogadro", Alessandria, Italy*

Abstract

Temporal constraints play a fundamental role in clinical guidelines. For example, temporal indeterminacy, constraints about duration, delays between actions and periodic repetitions of actions are essential in order to cope with clinical therapies. This paper proposes a computer-based approach in order to deal with temporal constraints in clinical guidelines. Specifically, it provides the possibility to represent such constraints and reason with them (i.e., perform inferences in the form of constraint propagation). We first propose a temporal representation formalism and two constraint propagation algorithms operating on it, and then we show how they can be exploited in order to provide clinical guideline systems with different temporal facilities. Our approach offers several advantages: for example, during the guideline acquisition phase, it enables to represent temporal constraints and to check their consistency; during the execution phase, it allows the physician to check the consistency between action execution-times and the constraints in the guidelines, and to provide query-answering and temporal simulation facilities (e.g., when choosing among alternative paths in a guideline).

Keywords:

artificial intelligence, clinical guidelines, temporal constraint representation, temporal reasoning, repeated/ periodic actions.

Introduction

Clinical guidelines are a means for specifying the "best" clinical procedures and for standardizing them. In recent years, the medical community has started to recognize that computer-based systems dealing with clinical guidelines provide relevant advantages, since, e.g., they can be used to support physicians in the diagnosis and treatment of diseases, or for education, critical review and evaluation aims [1]. Thus, many different approaches and projects have been developed to create domain-independent computer-assisted tools for managing clinical guidelines (see e.g., [2, 3]). Most of these approaches distinguish between an *acquisition phase*, in which expert-physicians (usually in cooperation with knowledge engineers) introduce clinical guidelines into the computer-based system, and an *execution phase*, when user-physicians execute a given guideline on a specific patient (i.e., on the basis of the patient's data). Moreover, recently, several approaches have started to focus also on the treatment of temporal aspects [4, 5, 6, 7]. As a matter of fact, in most therapies, actions have to be performed according to a set of temporal constraints concerning their relative order, their duration and the delays between them. Additionally, in many cases, actions must be repeated at regular (i.e., periodic) times. Furthermore, it is also necessary to carefully take into account the (implicit) temporal constraints derived from the hierarchical decomposition of actions into their components and from the control flow of actions in the guideline. A complete automatic treatment of temporal constraints involves – besides the design of an expressive representation formalism – also the development of suitable temporal reasoning algorithms operating on them, to be applied both at acquisition and at execution time. However, subtle issues such as the trade-off between the expressiveness of the representation formalism and the efficiency of correct and complete temporal reasoning algorithms have to be faced in order to deal with temporal constraints in a principled and well-founded way; few works in the area of computerized guidelines have deeply analyzed this topic so far.

In this paper, we first provide a brief overview of the state of the art and we discuss the advantages of a principled approach. Then, we introduce our representation formalism and our reasoning mechanisms to perform temporal reasoning in the acquisition and in the execution phase. Finally, we describe how to exploit our formalism and algorithms to provide clinical guidelines systems with temporal reasoning facilities, and we address comparisons and conclusions.

Background

State of the art

In the area of clinical guidelines, despite the large amount of work devoted to the *representation* of temporal constraints, little attention has been paid to temporal *reasoning*. Notable exceptions are represented by the approaches by Shahar [6] and by Duftschmid et al. [4]. In Shahar's approach, the goal of temporal reasoning is not to

deal with temporal constraints (e.g., to check their consistency), but to find out proper temporal abstractions to data and properties. Therefore, temporal reasoning is not based on constraint propagation techniques, in fact interpolation-based techniques and knowledge-based reasoning are used. Duftschmid et al. have proposed a comprehensive approach based on the notion of temporal constraint propagation [4]. In particular, in Duftschmid et al.'s approach, different types of temporal constraints – deriving from the scheduling constraints in the guideline, from the hierarchical decomposition of actions into their components and from the control flow of actions in the guideline – are supported. Temporal constraint propagation is used in order to (1) detect inconsistencies, and (2) provide the minimal constraints between actions. In [4], there is also the claim that (3) such a method can be used by the guideline interpreter to assemble feasible time intervals for the execution of each guideline activity.

New challenges and open problems

Despite the large amount of work, there still seems to be a gap between the range of phenomena covered by Artificial Intelligence temporal reasoning approaches and the needs arising from clinical guidelines management. In particular, in clinical guidelines the following features have to be supported:

1. qualitative and quantitative constraints, as well as repeated/periodic events; all types of constraints may be imprecise and/or partially defined;
2. a structured representation of complex events (in terms of part-of relations), to deal with structured descriptions of the domain knowledge;
3. the distinction between classes of actions (e.g. an action in a general guideline) and instances of such actions (e.g., the specific execution of an action in a guideline);
4. the consistency of the temporal constraints between classes and instances. This involves dealing with the inheritance of constraints (from classes to instances) and with the predictive role of constraints between classes[1].

Obviously, the interplay between issues (1)-(4) needs to be dealt with, too. For example, the interaction between composite and periodic events might be complex to represent and manage. In fact, in the case of a composite periodic event, the temporal pattern regards the components, which may, recursively, be composite and/or periodic events, as in Ex.1.

(Ex. 1) *The therapy for multiple myeloma is made by six cycles of 5-day treatment, each one followed by a delay of 23 days (for a total time of 24 weeks). Within each cycle of 5 days, 2 inner cycles can be distinguished: the melphalan treatment, to be provided twice a day, for each of the 5 days, and the prednisone treatment, to be provided once a day, for each of the 5 days. These two treatments must be performed in parallel.*

In Ex. 1, the instances of the melphalan treatment must respect the temporal pattern "twice a day, for 5 days", but such a pattern must be repeated for six cycles, each one followed by a delay of 23 days, since the melphalan treatment is part of the general therapy for multiple myeloma. Unfortunately, no current approach in Artificial Intelligence (henceforth AI) and in guideline literature proposes a comprehensive approach in which all the above phenomena can be represented, and correct, complete and tractable temporal reasoning can be performed. In this paper, we introduce an approach addressing all the above-mentioned issues.

Methods

Representing temporal constraints in clinical guidelines

Regarding temporal constraints concerning non repeated actions in the guidelines, we have chosen to model them using a well-known AI framework, STP (Simple Temporal Problem) [8]. This framework takes into account conjunctions (sets) of bounds on the distance between two time points (of the form $c \leq P1 - P2 \leq d$), and correct and complete temporal reasoning (e.g., for consistency checking) can be performed in cubic time by a classical all pairs shortest paths algorithm (such as Floyd-Warshall's one) [8]. STP allows to model precise and imprecise temporal locations (dates), durations, delays between points and qualitative temporal constraints (such as "A is before B").

Let us introduce first the constructs to model the repetition/periodicity constraints.

In our approach, the constraints on repetitions and periodicities are temporal constraints of the form

Repetition(A, RepSpec)

where *A* is a (possibly composite) action repeated according to the parameter *RepSpec*.

RepSpec is a recursive structure of arbitrary depth of the form

$$\mathbf{RepSpec = <R_1, R_2, \ldots, R_n>}$$

where each level R_i states that the actions described in the next level (i.e., R_{i+1}, or – by convention – the action *A*, if $i=n$) must be repeated a certain number of times in a certain time span. To be more specific, R_i consists of a quadruple

$$\mathbf{R_i = <nRepetitions_i, I\text{-}Time_i, repConstraints_i, conditions_i>}$$

where the first term represents the number of times that R_{i+1} must be repeated, the second one represents the time span in which the repetitions must be included, the third one may impose a pattern that the repetitions must follow, and the last one allows to express conditions that must

1 For example, given a guideline stating that action A_2 must be executed between one and two days after A_1, and given an execution of the action A_1 on a given patient at day d_1, one expects to have an instance of A_2 within day d_1+1 and day d_1+2.

hold so that the repetition can take place. Informally, we can roughly describe the semantics of a quadruple R_i as the natural language sentence "*repeat* R_{i+1} $nRepetitions_i$ *times in exactly* $I\text{-}Time_i$, *if* $conditions_i$ *hold*".

repConstraints$_i$ is a (possibly empty) set of pattern constraints, representing possibly imprecise repetition patterns. Pattern constraints may be of type:

- **fromStart(min, max)**, representing a delay between the start of the I-Time and the beginning of the first repetition;
- **toEnd(min, max)**, representing a delay between the end of the last repetition and the end of the I-Time;
- **inBetweenAll(min, max)**, representing the delay between the end of each repetition and the start of the subsequent one;
- **inBetween((min_1,max_1),…,($min_{nRepetitions_{i-1}}$,$max_{n\text{-}Repetitions_{i-1}}$))** representing the delays between each repetition and the subsequent one. Note that any couple (min_j, max_j) may be missing, to indicate that the constraint does not impose any temporal constraint between the j^{th} repetition and the $(j+1)^{th}$ one.

Let us see an example to illustrate the use of *repConstraints*:

(Ex. 2) Intrathecal methotrexate must be administered 7 times during 88 weeks, never less than 10 weeks apart or more then 14 weeks apart.

Ex. 2 may be represented with a one-level specification:

Repetition(Intrathecal_methotrexate,
< <7,88wk,{inBetweenAll(10wk, 14wk)}, ∅>>).

It is worth noting that $repConstraints_i$, $nRepetitions_i$ and $conditions_i$ are not mandatory.

conditions$_i$ is a (possibly empty) set of conditions that influence the repetitions. The conditions may be of type:

- **while(B)**, where B is a Boolean predicate. It states that, as soon as B becomes false, a break from the repetitions is forced, i.e., the repetitions must immediately interrupt. As an example, we may consider the following:

(Ex. 3) Give acetaminophen twice a day until the fever has gone.

This may be represented as:

Repetition(acetaminophen,
< <_, _, ∅, {while(fever)}>, <2, 1d, ∅, ∅> >.

- **onlyIf(B)**, where B is a Boolean predicate. It states that, if B is true, the repetition may be performed and, if B is false, the repetition must not be performed and we can pass to the next repetition. This construct allows to skip single repetitions. As an example, we may consider the following:

(Ex. 4) Give acetylsalicylic acid twice a day for a maximum of 15 days, only if there is migraine.

This may be represented as:

Repetition(acetylsalicylic_acid,
<<_, 15d, ∅, ∅>, <2, 1, ∅, onlyIf(Migraine)>>.

The formalism we are introducing allows one to manage different kinds of imprecision; in fact:

- there may be arbitrary delays between the repetitions;
- the *(min, max)* specifications in $repConstraints_i$ make it possible to specify variable delays between the repetitions.

Dealing with imprecise temporal constraints is very important for the practical applicability of the approach.

Moreover, the repetitions may be nested at arbitrary depth, representing simple cases with fewer levels as in Ex. 2 and more complex cases with more levels as in Ex. 5, an excerpt from a clinical guideline for the treatment of Childhood Acute Lymphoblastic Leukaemia:

(Ex. 5) The therapy lasts 88 weeks and it is repeated twice in four years. In the therapy, cotrimoxazole must be given twice daily on two consecutive days every week.

Ex. 5 may be represented in the following way:

Repetition(Cotrimoxazole, < <2, 4y, ∅, ∅>, <_, 88wk, ∅, ∅>, <2, 1wk, {inBetweenAll(0,0)}, ∅>, <2, 1d, ∅, ∅>>), where the pattern constraint *inBetweenAll(0,0)* in the third triple imposes that the days must be consecutive.

In order to make our approach to temporal constraints more user-friendly, a (possibly graphical) interface could be used to acquire and represent temporal constraints (concerning both (i) dates, durations, delays and qualitative relations between non-repeated events and (ii) repetition/ periodicity constraints).

Reasoning with temporal constraints in clinical guidelines

Internal representation

STP provides a suitable representation for temporal constraints on non-repeated actions. In fact, the constraints can be modeled as graphs on which the well-known Floyd-Warshall's algorithm operates to check consistency. However, the STP framework is not expressive enough to cope with repeated/periodic actions. Thus, we have chosen to model the constraints regarding repeated actions into separate STPs, one for each repeated action. Thus, in our approach, the overall set of constraints between actions in the guideline is represented by a tree of STPs (STP-tree henceforth). The root of the tree (node N1 in the example in figure 1) is the STP which homogeneously represents the constraints (including the ones derived from the control flow of actions in the guideline) between all the actions in the guideline (e.g., in N1, the fact that the duration of the chemotherapy is 168 days), except repeated actions. Each node in the tree is an STP, and has as many children as the number of repeated actions it contains. Each edge in the tree connects a pair of endpoints in an STP (the starting and ending point of a repeated action) to the STP containing the constraints between its subactions and it is labeled with the list of properties describing the temporal constraints on the repetitions (i.e., *RepSpec*). For

example, in Fig. 1, we show the STP-tree representing the temporal constraints involved by Ex. 1.

Additionally, an independent STP must be used in order to represent the temporal constraints about the specific *instances* of the actions of the guidelines, as emerging from executions of the guidelines on specific patients.

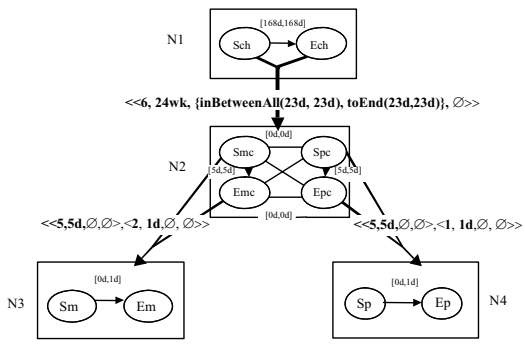


Figure 1 - STP-tree for the multiple myeloma chemotherapy guideline in Ex. 1. Thin lines and arcs between nodes in a STP represent bound on differences constraints. Arcs from a pair of nodes to a child STP represent repetitions. Arcs between any two nodes X and Y in a STP of the STP-tree are labeled by a pair [n,m] representing the minimum and maximum distance between X and Y. Sch, Ech, Smc, Emc, Spc, Epc, Sm, Em, Sp and Ep stand for the starting (S) and ending (E) points of chemotherapy, melphalan cycle, prednisone cycle, melphalan treatment and prednisone treatment, respectively.

Checking the consistency of a guideline

In order to check the consistency of the STP-tree, it is not sufficient to check the consistency of each node separately. In such a case, in fact, we would neglect the repetition/periodicity information. Temporal consistency checking, thus, proceeds in a top-down fashion, starting from the root of the STP-tree. Basically, the root contains a "standard" STP, so that the Floyd-Warshall's algorithm can be applied to check its consistency. Thereafter, we proceed top down towards the leaves of the tree. For each node *X* in the STP-tree (except the root), we progress as in the *STP_tree_consistency* algorithm:

1. temporal constraints on the incoming arcs are inherited by X;
2. STP+inherited temporal constraints are propagated by means of Floyd-Warshall's algorithm;
3. check that the propagated temporal constraints are consistent.

Property. The *STP_tree_consistency* algorithm is correct, complete and tractable.

Reasoning with the executions of the guideline

In the following, we report an algorithm for checking the consistency of the execution of a guideline with respect to its related guideline. In our work, as in most approaches to clinical guidelines, we suppose that one has *full observability* of instances (i.e., all the instances of actions which have been executed have been observed and inserted into the knowledge base), and that, for each instance, one knows the corresponding class of actions and/or repetition in the guidelines.

Algorithm *integratedConsistency*:

1. check that all and only the instances predicted by the guideline are present;
2. inherit the repetition/periodicity constraints and the temporal (non-periodic) constraints from the guideline to the instances;
3. propagate the temporal constraints via the Floyd-Warshall's algorithm;
4. check that the propagated temporal constraints are consistent.

Property. The *integratedConsistency* algorithm is correct, complete and tractable.

Architecture of GLARE

In order to enhance the generality of the temporal reasoning approach, the temporal reasoning algorithms are provided by a modular approach, in which a layered Temporal Server (TS) is loosely coupled with a guideline system (see figure 2). The clinical guideline system delegates temporal-related problems to the TS module. The core of TS is the *temporal reasoner* (TR), which implements the temporal reasoning algorithms and the related data structures. The *facilities layer* uses the two consistency-checking algorithms to provide the facilities (a)-(d) described below. Moreover, for acquiring and representing temporal information the *interface layer* may make use of visualization techniques.

Results

Exploiting temporal reasoning in clinical guidelines systems

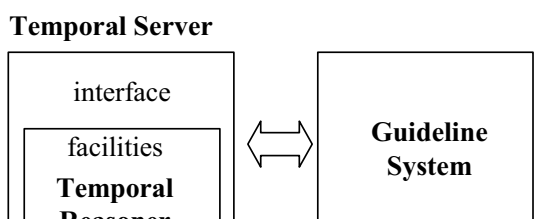


Figure 2 - Loosely coupled architecture to extend a guideline system with temporal reasoning facilities

In the previous sections, we have proposed a principled approach coping with issues (1)-(4). The adoption of our approach can provide computer-based guideline systems with crucial advances. In the following, we discuss several facilities that can be designed on the basis of our representation formalism and reasoning algorithms, both during guideline acquisition and execution. Although the

approach we propose is system-independent, in some cases we will exemplify it by sketching how we are planning to implement it in GLARE (GuideLine Acquisition, Representation and Execution) [9, 10, 11]. GLARE is a prototypical system to acquire and execute clinical guidelines, developed by the Computer Science Department of the Università del Piemonte Orientale of Alessandria (Italy) in cooperation with Azienda Ospedaliera San Giovanni Battista of Torino (the second hospital in Italy).

During **acquisition**, expert physicians (usually with the help of a knowledge engineer) represent a clinical guideline in a computer-based formalism, considering also the temporal constraints it contains. An automated ***consistency-checking*** facility is required to grant the temporal consistency of the guideline in a principled way. Such a facility can be provided through a call to the *STP_tree_consistency* algorithm, which can be advocated at any stage during the acquisition of a clinical guideline, so that incremental consistency checking is also possible. By default, consistency checking can also be executed at the end of each acquisition working session.

During **execution**, a guideline (e.g., the guideline for asthma) is applied to a specific patient, i.e., specific instances of the classes of actions in the guideline are executed. Our *integratedConsistency* algorithm can be exploited as the core of a user-oriented approach providing some crucial temporal facilities to user-physicians. Several of these facilities rely on the fact that our algorithm also provides the *minimal network* of the constraints between instances (considering also the constraints inherited from the guideline).

a) First of all, for scheduling purposes, it is important to provide a facility to assess when the next actions have to be performed, given the constraints in the whole guideline and given the time when the last actions in the guideline have been executed. The execution time of the next action(s) can be obtained through the steps 1-3:

1. retrieve the set of candidate next actions through a navigation of the control flow relations in the guideline;
2. apply the *integratedConsistency* algorithm to obtain the minimal network of temporal constraints;
3. retrieve the actions' possible execution-times from the minimal network (in the form of distances from the last-executed action, or from the origin of time).

By combining this facility with the query answering one (see facility (c) below), temporal reasoning can also be used in an interactive way to determine schedules which are consistent with the temporal constraints. For example, given a pattern $A_1, \ldots, A_n$ of actions in a guideline, temporal reasoning can be used in order to answer queries such as "*If I perform action A_1 today at 11 am, when will I have to perform $A_2, \ldots, A_n$?*", or "*Is it OK if I perform A_1 today at 12 am, A_2 at 6 pm and A_3 at 8 pm, and, if so, when will I have to perform A_4?*";

b) From the point of view of quality evaluation/assessment, it is important to provide a facility to check whether the temporal constraints in the guideline have been respected or not by the instances of actions that have been executed (considering also partial – i.e., ongoing – executions). Such a facility is directly provided by the *integratedConsistency* algorithm;

c) One of the main goals of guideline computer-based systems is to support decision making. In such a context, providing a (temporal) query-answering facility is a crucial task. Such a facility can be efficiently implemented on the basis of the minimal network provided by the *integratedConsistency* algorithm (along the lines discussed in [12]0), both to answer yes/no queries (e.g., "*is it correct to execute action A now, and action B within the next two hours?*") and/or to have in output the minimal distance between the instances of actions;

d) Still considering decision making, temporal reasoning can be profitably coupled with "simulation" computer-based facilities to see the temporal consequences of choosing among different alternative paths in a guideline. In particular, GLARE provides the "what if?" facility allowing physicians to discriminate among different alternatives of a decision by simulating the consequences of each choice, i.e., by visiting the paths in the guideline stemming from each one of the alternatives (see, e.g., [10]). Taking advantage of the *integratedConsistency* algorithm, our approach provides physicians with a way of comparing paths from the temporal point of view (i.e., in order to find the maximal and minimal temporal duration of each path). This facility can be obtained as follows:

1. for each path P_i to be compared
2. hypothesize the existence of an instance of each action in P_i which has not yet been executed;
3. apply the *integratedConsistency* algorithm;
4. retrieve the minimal and maximal duration of P_i.

Discussion and conclusion

In this paper, we propose a principled domain- and system-independent approach to the treatment of temporal constraints in clinical guidelines. We introduce a new representation formalism, coping with different types of temporal constraints, including constraints on (possibly periodic) repeated events. We devise two reasoning mechanisms to check the consistency of temporal constraints in clinical guidelines. Finally, we show how they can be exploited to provide clinical guideline systems with temporal reasoning facilities.

The approach by Duftschmid et al. [4]0 is the closest one to ours in the literature. With respect to such an approach, we propose an extended language to deal with repetitions (e.g., we cope with conditioned repetitions, through the '*while'* and '*onlyIf'* constructs) and we extend the basic STP framework, via the definition of the STP-tree and of the related constraint propagation algorithms. Finally, from the point of view of end-users, we also provide the (a)–(d) facilities discussed in the previous section. In particular, the treatment of the "a-posteriori" consistency between the temporal constraints in the guideline and of

the execution times requires several extensions both from the representation point of view (since a separate STP needs to be used) and from the algorithmic point of view (since new constraint-propagation-based temporal reasoning algorithms have to be devised).

Acknowledgments

We are very grateful to Prof. Gianpaolo Molino, Azienda Ospedaliera San Giovanni Battista, Turin, Italy and his group for their continuous advise and cooperation in developing and testing the GLARE system, and, specifically, the temporal approach described in this paper.

References

[1] Gordon C and Christensen JP, eds., Health Telematics for Clinical Guidelines and Protocols (IOS Press, Amsterdam), 1995.

[2] JAMIA, Focus on Clinical Guidelines and Patient Preferences, JAMIA 15(3), 1998.

[3] Peleg M, Tu S, Bury J, Ciccarese P, Fox J, Greenes RA, Hall R, Johnson PD, Jones N, Kumar A, Miksch S, Quaglini S, Seyfang A, Shortliffe EH, Stefanelli M., Comparing computer-interpretable guideline models: a case-study approach, J Am Med Inform Assoc.,10(1):52-68, 2003.

[4] Duftschmid G, Miksch S, Gall W, Verification of temporal scheduling constraints in clinical practice guidelines, Artificial Intelligence in Medicine 25(2): 93-121, 2002.

[5] Keravnou ET, Special issue: Temporal Reasoning in Medicine, Artificial Intelligence 8(3), 1996.

[6] Shahar Y, A Framework for Knowledge-Based Temporal Abstraction in Clinical Domains, Artificial Intelligence 90(1):79-133, 1997.

[7] Shahar Y, Musen MA, Knowledge-based temporal abstraction in clinical domains, Artificial Intelligence in Medicine 8(3):267-298, 1996.

[8] Dechter R, Meiri I, Pearl J, Temporal Constraint Networks, Artificial Intelligence 49:61-95, 1991.

[9] Terenziani P, Molino G, Torchio M, A Modular Approach for Representing and Executing Clinical Guidelines, Artificial Intelligence in Medicine 23:249-276, 2001.

[10] Terenziani P, Montani S, Bottrighi A, Molino G, Torchio M, Supporting physicians in taking decisions in Clinical Guidelines: the GLARE's "what if" facility, Proc. AMIA 2002, 772-776, 2002.

[11] Terenziani P, Montani S, Torchio M, Molino G, and Anselma L, Temporal Consistency Checking in Clinical Guidelines Acquisition and Execution: the GLARE's Approach, JAMIA, Special issue on the Fall AMIA Symposium, pp. 659-663, 2003.

[12] Brusoni V, Console L, Pernici B, Terenziani P, LaTeR: managing temporal information efficiently, IEEE Expert 12(4):56-64, 1997.

Address for correspondence

Address: Corso Svizzera 185, 10149 Torino, Italy
E-mail: anselma@di.unito.it

Chapter 6.

Improving Quality

This page intentionally left blank

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Information and Communication Processes in the Microbiology Laboratory – Implications for Computerised Provider Order Entry

Andrew Georgiou[a], Joanne Callen[b], Johanna Westbrook[a], Mirela Prgomet[b], George Toouli[c]

[a] *Health Informatics Research & Evaluation Unit, University of Sydney*
[b] *Faculty of Health Sciences, University of Sydney*
[c] *South Western Area Pathology Services, Liverpool, Sydney*

Abstract

The aim of this multi-method study based at a microbiology department in a major Sydney metropolitan teaching hospital was to: i) identify the role that information and communication processes play in a paper-based test request system, and ii) examine how these processes may affect the implementation and design of Computerised Provider Order Entry (CPOE) systems. Participants in this study reported that clinical information can impact on the urgency and type of tests undertaken and affect the interpretation of test results. An audit of 1051 microbiology test request forms collected over a three-day period showed that 47% of request forms included clinical notes which provide a variety of information often vital to the test analysis and reporting process. This transfer of information plays an important role in the communication relationship between the ward and the laboratory. The introduction of new CPOE systems can help to increase the efficiency of this process but for that to be achieved research attention needs to be given to enhancing the provision and communication of clinical information.

Keywords:

Evaluation studies, hospital information systems, laboratories, microbiology, pathology, qualitative research

Introduction

Computerised Provider Order Entry (CPOE) systems pose major challenges for hospital pathology laboratories [1], with important implications for a range of laboratory processes including inter-department functions, work organisation and laboratory effectiveness [2-5]. CPOE systems provide clinicians with the ability to enter orders directly into a computer [6]. The incorporation of functions such as clinical decision support and patient database linkage provide the potential to significantly impact on the quality of health care delivery leading to improved patient outcomes [7-9].

However, within the research and medical literature there has been relatively little attention given to the effect of CPOE on pathology laboratories [10]. These services play a crucial role in overall patient safety and outcome, accounting for an estimated 70% of all information used in decision making for admission, treatment and discharge [11, 12]. Pathology services are information intense units reliant on the efficient management and timely communication of relevant information to maximize the delivery of health care [13]. Moreover, the pathology department is comprised of a number of organisational structures and bodies each with its own unique and highly complex way of performing tasks and interacting with other departments. The aim of this study was to: i) identify the role that information and communication processes play in a paper-based test request system in the microbiology department; and ii) examine how these processes may affect the implementation and design of CPOE systems.

Methods

Design and research setting

A multi-method study (using quantitative and qualitative data collection techniques) was conducted in a microbiology department based at an Australian metropolitan teaching hospital. The department receives 131,000 microbiology test requests and specimens annually and employs 55 staff. It is divided into bacteriology, molecular biology, serology, virology, mycobacteriology and parasitology sections. The department is part of a pathology service involved in the introduction of Cerner Millennium PowerChart (Version 2004.01). The pathology service is responsible for a large metropolitan area and provides diagnostic services to a number of hospitals (including teaching hospitals) and clinics. This study forms part of a large multi-hospital research project evaluating the implementation of CPOE with data collection occurring from sites pre- and post- implementation.

Qualitative data relating to existing information and communication processes connected with test ordering and reporting within the microbiology laboratory were collected by observations, interviews and a focus group. Quantitative data collected from the microbiology department consisted of the volume of tests ordered and measured the presence of additional clinical information provided by doctors. The results provided a baseline indication of the existence of clinical information on test request forms by the requesting doctor.

Selection and sampling logic

The site was chosen because it was about to introduce an electronic ordering system that was mandatory for all inpatients. Qualitative data collection began with a focus group consisting of four laboratory scientists and one laboratory manager (n=5). These participants were chosen for their suitability (i.e., the department manager attested to their experience and knowledge of microbiology laboratory processes). The aim of the focus group was to discuss participants' views and expectations of the impact of a new CPOE system that was due to be introduced within the next three months. A set of semi-structured questions were used to gather impressions about the current system of paper-based test requests and what changes participants thought the new system would introduce. Participants were asked to raise both positive and negative features of the current laboratory processes. The participant base was extended using purposive sampling techniques [14], whereby participants directed us to other key informants. This increased the number of participants to eleven. Overall it included two senior laboratory scientists, one laboratory business manager, three technical officers and five laboratory scientists. Interviews were repeated with participants for clarification and further exploration of issues raised, with a total of 20 interview sessions conducted. This process provided a valuable feedback mechanism which enhanced our confidence in the validity of emerging themes [15]. One researcher conducted all the interviews and the focus group. Eight observation sessions were conducted by two researchers with each session lasting on average 1.5 hours (total of 12 hours of observations). One researcher conducted five of the observation sessions and the other conducted the remaining three.

Data collection

In the course of our analysis of the qualitative data, the research team undertook the collection of microbiology test request data. This provided the study with an important triangulation technique to investigate emergent themes [16]. Request forms were audited by one researcher over a three-day period between 29 June 2006 and 1 July 2006. No details related to patient identification were collected. The data collected included the number of test request forms received with and without the inclusion of clinical notes. For the purposes of this study clinical notes refers to patient specific clinical information written on the laboratory request form by the doctor requesting the test. Clinical notes therefore, may include: signs and symptoms; the site from which the specimen was obtained; medical history; physical examination; medications and the reason for the test request. One of the above pieces of clinical information was needed for the test request form to satisfy the criteria of "test request contained clinical notes."

The observations, interviews and focus group were conducted between August 2005 and October 2006. A letter outlining the study, its voluntary nature, the confidentiality of findings and participants, and a consent form, were provided to all participants. The research was approved by the University of New South Wales Human Research Ethics Committee and the relevant Area Health Service Research Ethics Committee.

Data analysis

The quantitative data were entered into SPSS (Version 12.0.1 for Windows 2004) and analysed using descriptive statistics. The focus group and one interview were recorded and transcribed. The remaining interviews and observations were recorded by the researchers in note form. The qualitative data were interpreted using a grounded theory approach [16] to derive themes that explained the information and communication processes within the microbiology department. Triangulation of analysis involved a number of iterative sessions involving a total of five researchers discussing and analysing the data: two who had collected the data and three others from the research team [14].

Results

The results are presented in two parts: firstly the qualitative data about laboratory information and communication processes related to the test ordering process, and secondly the volume, type and inclusion of clinical notes on microbiology test requests.

Laboratory information and communication processes related to the test ordering process – qualitative analysis

Three themes relating to information and communication processes surrounding test requests were identified:

- Theme 1: The context of the microbiology department
- Theme 2: Communication of clinical information
- Theme 3: Expectations of the new electronic ordering system

The context of the microbiology department

Participants explained that microbiology departments have their own specific requirements and needs that are not always applicable to other departments. For instance, the issue of timeliness has a particular context-dependent meaning for microbiology that is not identical to other pathology departments (e.g. biochemistry) for whom the optimisation of turnaround times for the processing and issue of results is an important performance indicator. Microbiology deals predominantly with diseases caused by infectious agents (e.g. bacteria, viruses, fungi and parasites) requiring time to grow before an appropriate test result is available.

The communication of clinical information

Participants highlighted the provision of relevant and appropriate clinical information by doctors ordering tests as a key area that directly impacts on their efficiency. In hospitals where electronic ordering has not been implemented this means the provision of a hand written test request form, including clinical notes, from the requesting doctor. If clinical information is not included the request may be judged to be incomplete or inadequate and in need of some form of follow up, often through direct telephone contact with the requesting doctor. This point was described by one participant in the following way:

> "As a whole the request that we receive, we need to know what the specimen is. We need to know what they want us to do with it, and it needs to be legible, so it really is an error, because we have to use our time to verify what they actually want." (Focus group participant)

Clinical notes are very important to laboratory staff. This is because they play an important role in setting the context for the test. Laboratory managers explained that this contextual information improves the laboratory's input. For instance, it may help a pathologist detect the need for more tests, or perhaps identify when a doctor may have asked for an inappropriate test.

> "They don't tell us what they want and we process what we think. If we didn't get the correct clinical details we may not necessarily make it up for the right thing…. (Focus group participant)

A salient example of this is for the disease tuberculosis, which the laboratory may not routinely test for unless it is either specifically requested, or when relevant clinical information is provided.

> "There are times when we process a specimen, then they [clinicians] ring up and say: have you done TB [tuberculosis] on this? We say – well you didn't ask for it. They should have given us the clinical details that would have allowed us to do that." (Focus group participant)

Expectations of new electronic ordering

The introduction of electronic ordering was expected to alter the way the department communicates with clinicians. In particular, laboratory personnel would not be required to decipher hand written notes anymore, which should eliminate instances of unclear or illegible requests. Participants described the potential of more effective exchange of valuable and relevant clinical information.

> "There should be some benefits to the laboratory, in that there will be less data entry, I guess. The patients' demographics etc, will come across. There should be less confusion, as to what tests are requested by the medical staff. We are hoping to get a lot more clinical details…" (Focus group participant)

The volume and inclusion of clinical notes on test request forms – quantitative analysis

The total number of microbiology specific tests requested within the 1,051 test request forms received were 1,078 as some request forms contained multiple test requests (Table 1). A large proportion of test request forms (47%) contained clinical notes documented by the clinician on the request form. The average number of tests requested per day, over the 3 day period, was 359 (range 338 to 374).

Table 1 – Number of microbiology specific tests requested with and without clinical notes

	No of tests (n=1078)	**With Clinical notes (n=506)**	
Day	n	n	%
1	374	186	49.7
2	338	146	43.2
3	366	174	46.9
Average	359.3	168.7	46.9

Table 2 highlights the results of the most frequently ordered tests. The most requested tests were urine cultures (35%) followed by blood cultures (21%) and specific site swab cultures (8%). The majority of urine culture requests (n=233[62%]) and blood culture requests (n= 142[62%]) did not contain clinical notes. However most wound culture requests (n=42[75%]) did contain clinical notes.

Table 2 – Frequency of a selection of the most ordered microbiology requested tests

Test categories	**No of tests (n=1078)**	**%**
Non-specific site swab cultures	27	3
Stool cultures	35	3
Infection control cultures	38	4
Fluid cultures	44	4
Sputum cultures	52	5
Wound cultures	56	5
Genital cultures	74	7
Specific site swab cultures	81	8
Blood cultures	230	21
Urine cultures	379	35
Others	62	6

Discussion

The comparison of results collected from the audit of microbiology test requests with the themes identified from the interviews and focus group session provides a means to triangulate different types of data, and encourages a better understanding of the meaning and significance of different findings. The results showed an important proportion (47%) of microbiology test requests received by the microbiology department contained some clinical notes

provided by clinicians. This indicates that clinicians often need to communicate further information to the microbiology department beyond simply identifying the test to be performed. There are cases, as in most blood and urine cultures, (which make up the bulk of tests requested), that do not contain any clinical notes. In some instances, (as suggested in the interviews and focus groups) this may require laboratory staff to follow up the missing information using telephone communication.

The translation of data into clinically meaningful information

The results highlight the role that the supply and processing of clinical information plays in the microbiology laboratory. The traditional format through which this information is communicated has been the hand written request form. Aside from their obvious clerical function, these forms also represent an important link between doctors and the laboratory [17] through which contextual patient data are communicated. This information can impact on the urgency, choice and even interpretation of pathology tests and results. The laboratory process involves the translation of data into clinically meaningful information. This role can be described as a core function of the laboratory service [18] and represents an important contribution to the patient care process [19]. For some commentators, such as Marques and MacDonald the absence of clinical information in certain situations can be misleading and even potentially dangerous [20].

Communication and the transfer of information

This study also demonstrates that the exchange of information across the hospital is a two-way process through which clinicians not only provide clinical information to laboratories, but also receive it back in an enhanced form. This relationship demonstrates the importance that communication plays in this process. The ordering process can be conceptualised as part of a collaborative effort of multidisciplinary groups [21]. For Toussaint and Coiera every information exchange is a communication act including a simple exchange between two people or even two machines [22]. Communication systems are important parts of the information structure [23] and can play a major role in the decision making process.

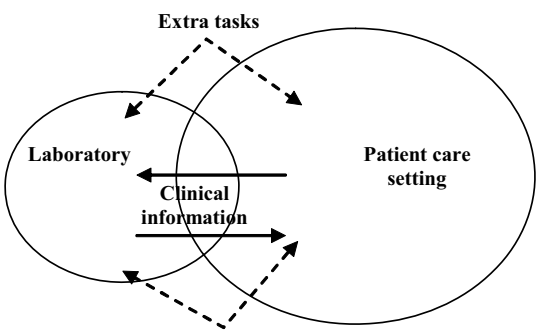


Figure 2 – The laboratory/ward information and communication relationship

Most information transactions within health services occur without the involvement of electronic data or systems [23], usually in conversations or through paper exchange. In reality, hospital communication systems form a collection of differing components and types [24]. Electronic systems like CPOE will impact significantly on existing channels and relationships [2]. The results of this study suggest that CPOE systems can be expected to enhance communication ensuring legibility and clarity in the ordering process and contribute to improvements in the clinical decision making process [4, 7]. However, there is also evidence that CPOE systems may disrupt previous channels of communication and replace them with inadequate or unsuitable alternatives often involving workarounds and extra tasks [25]. As Gorman et al. have stressed, incomplete or incorrect models of the process can lead to problems in the uptake and operation of CPOE [21]. Figure 2 depicts the importance of clinical information for the communication exchange between the laboratory and the clinical setting. The broken lines highlight the potential for this flow to be disturbed by design inconsistencies and barriers. It is important therefore that information and communication processes (at times unique to each hospital) are clearly identified as a means of maximizing the benefits of CPOE systems.

Limitations of this study

This study was undertaken at one site, using a microbiology laboratory department during the lead up to the implementation of a new CPOE system. The experiences of this one site will not be identical to other laboratory departments in other hospitals. Nevertheless the issues outlined will have wider resonance. The multi-method design adopted by this study has the advantage of providing rich contextual qualitative data about how the department's information and communication requirements are perceived along with descriptive data summarising the existing arrangements using hand written requests. This multi-method approach helped to enhance the findings and inform the discussion with participants. The results provide a useful evaluative framework with which to approach the question of CPOE implementation. But it also suggests the need to closely examine and quantify the impact different types of clinical information provided for different test requests can have on the laboratory process and their subsequent communication with doctors. While this task was beyond the scope of this study, it remains a natural area for follow up.

Conclusion

This study underscores the important role that the provision, processing and exchange of clinical information plays in microbiology laboratory processes. Clinical information helps to inform the laboratory of the type and urgency of tests required as well as assisting pathology staff to add interpretative value to the information provided back to medical staff. The exchange and transfer of clinical information is underpinned by a complex variety of communication channels within the hospital. New CPOE systems can increase the efficiency of this process and enhance the richness of the information exchange. To date, little attention has been provided to this issue. We recommend that more research into this area be undertaken so as to make these channels of communication and information exchange more explicit, and as a means of

providing information to enhance the design and implementation of CPOE systems.

Acknowledgements

This study was part of an Australian Research Council Linkage Grant project that funded the evaluation of the impact of information and communication technologies on organisational processes and outcomes. The project was undertaken with the collaboration of the New South Wales Health Department. The authors also acknowledge laboratory personnel for their participation in this study.

References

[1] Davidson EJ and Chismar W. Examining the organizational implications of IT use in hospital-based health care: a case study of computerized order entry. in Proceedings of the 32nd Hawaii International Conference on System Sciences. 1999. Hawaii.

[2] Stablein D, Welebob E, Johnson E, Metzger J, Burgess R, and Classen DC, Understanding hospital readiness for computerized physician order entry. Jt. Comm J Qual Saf 2003;29(7); 336-44.

[3] Massaro TA, Introducing physician order entry at a major academic medical center: I. Impact on organizational culture and behavior. Acad Med 1993;68(1); 20-25.

[4] Mekhjian HS, Kumar RR, Kuehn L, Bentley TD, Teater P, Thomas A, Payne B, and Ahmad A, Immediate Benefits Realized Following Implementation of Physician Order Entry at an Academic Medical Center. J Am Med Inform Assoc 2002;9(5); 529-539.

[5] Kuperman GJ, Boyle D, Jha A, Rittenberg E, Ma'Luf N, Tanasijevic MJ, Teich JM, Winkelman J, and Bates DW, How Promptly Are Inpatients Treated for Critical Laboratory Results? J Am Med Inform Assoc 1998;5(1); 112-119.

[6] Doolan DF and Bates DW, Computerized physician order entry systems in hospitals: mandates and incentives. Health Aff 2002;21(4); 180-8.

[7] Sittig D and Stead W, Computer-based physician order entry: the state of the art. J Am Med Inform Assoc 1994;1(2); 108-123.

[8] Abelson D, Connelly D, Klee G, Maag T, and Smith J, Technology Assessment: Computerized Physician Order Entry. 2001, Institute for Clinical Systems Improvement.

[9] The Leapfrog Group for Patient Safety, Fact Sheet: Computerized Physician Order Entry (CPOE) (2003) Available at: http://www.leapfroggroup.org/media/file/Leapfrog-Computer_Physician_Order_Entry_Fact_Sheet.pdf Accessed 2 March 2006.

[10] Georgiou A, Williamson M, Westbrook J, and Ray S, The impact of computerised physician order entry systems on pathology services: a systematic review (doi:10.1016/j.imedinf.2006.02.04). (In press) Int J Med Inf 2007.

[11] Forsman RW, Why is the laboratory an afterthought for managed care organizations? Clin Chem 1996;42(5); 813-6.

[12] Becich MJ, Information management: moving from test results to clinical information. Clin Leadersh Manag Rev 2000;14(6); 296-300.

[13] Review of NHS Pathology Services in England, Report of the Review of NHS Pathology Services in England 2006 Available at: www.dh.gov.uk Accessed 10 August 2006.

[14] Bowling A, Research methods in health. Open University Press: Buckingham; 1997.

[15] Guba E and Lincoln Y, Fourth generation evaluation. Sage: Newbury Park, CA; 1989.

[16] Glaser B and Strauss A, The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine Publishing Company: New York; 1967.

[17] Deeble J and Lewis-Hughes P, Directions for pathology. National Health Strategy Background Paper No 6. National Health Strategy: Melbourne; 1991.

[18] Travers E, Management of Information and its Economic Impact in Clinical Laboratory Management, E. Travers, Editor. 1997, Williams & Wilkins: Baltimore.

[19] Hardwick DF, "Directoring" and managing in a professional system. Modern Pathology 1998;11(6); 585-92.

[20] Marques MB and McDonald JM, Defining/measuring the value of clinical information. Clin Leadersh Manag Rev 2000;14(6); 275-9.

[21] Gorman PN, Lavelle MB, and Ash JS, Order creation and communication in healthcare. Methods Inf Med 2003;42(4); 376-84.

[22] Toussaint PJ and Coiera E, Supporting communication in health care. Int J Med Inf 2005;74(10); 779-81.

[23] Coiera E, When conversation is better than computation. J Am Med Inform Assoc 2000;7(3); 277-86.

[24] Coiera E, Communication Systems in Healthcare. The Clinical Biochemist Reviews 2006;27; 89-98.

[25] Georgiou A, Westbrook J, Braithwaite J, Iedema R, Ray S, Forsyth R, Dimos A, and Germanos T, When requests become orders - A formative investigation into the impact of a computerized physician order entry system on a pathology laboratory service (In press). (doi:10.1016/j.medin.2006.04.0020) Int J Med Inf 2007.

Address for correspondence

Health Informatics Research & Evaluation Unit
Cumberland Campus, University of Sydney,
PO Box 170 Lidcombe NSW 1825, Australia
Andrew Georgiou (Senior Research Fellow)
Telephone: + 61 2 9036 7331 Email: a.georgiou@usyd.edu.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Using an Accident Model to Design Safe Electronic Medication Management Systems

Farah Magrabi[1], Geoff McDonnell[1], Johanna I. Westbrook[2], Enrico Coiera[1]

[1]Centre for Health Informatics, University of New South Wales, Australia
[2]Health Informatics Research & Evaluation Unit, Faculty of Health Sciences, University of Sydney

Abstract

Large-scale implementation of electronic prescribing systems (e-PS) is likely to introduce at least some machine-related errors that will harm patients. We present a dynamic systems modeling approach to developing a comprehensive multilevel accident model of the process, context and task interaction variables which give rise to human error and system failure when e-PS are used in routine care. System dynamics methods are used to represent interactions between medication management processes and the context that is relevant to error generation, interception and transmission, agent-based methods represent task interactions. Capturing the patterns of failure within an accident model will facilitate an evidence-based approach to hazard analysis and design of e-PS features to improve patient safety. The model will have broad potential to guide the design, implementation and regulation of e-PS.

Keywords:

safety management, medical order entry systems, medication errors, computer simulation, systems theory.

Introduction

Medication errors are one of the most significant causes of iatrogenic harm and death within health care systems internationally [1]. Electronic medication management systems or prescribing systems (e-PS) are strongly recommended by health care bodies as an effective patient safety intervention to reduce medication errors [1-3]. Because of the widely reported potential benefits of e-PS in improving patient safety, quality of care and efficiency in healthcare delivery, these systems are central to health information and communications technology (ICT) implementation strategies worldwide [2, 3].

Safety of e-Prescribing Systems – an emerging problem in healthcare

Alongside their potential benefits, ICT systems can also be a source of harm to patients [4], and this has sparked urgent debate in the international Health Informatics community about the risks of harm associated with introducing this new technology. Several recent studies have identified the potential of these systems to have an adverse impact on patient safety [5]. For example, one study generating much controversy following its publication in December 2005 reported a significant increase in patient deaths following the introduction of electronic orders at the University of Pittsburgh Children's hospital (mortality rate increased from 2.80% to 6.57% post-implementation) [6, 7]. In another US study, Koppel et. al. found that computerization within a Pennsylvania hospital facilitated 22 types of medication errors [8]. The US Pharmacopoeia, a drug industry standards group which monitors patient safety, reported a steady growth in medication errors associated with e-PS [9]. Nearly 20% of these errors reported in 2003 were automation related.

Need for safety interventions

Evaluations of e-PS are increasingly identifying the need for patient safety interventions to minimize the impact of human error and system failure associated with ICT use in routine clinical tasks. For example, an evaluation of an e-PS within a Veterans Administration Medical Centre in Salt Lake City found that 93% of adverse events could be prevented through the implementation of a safety feature that would individualize safe dosage information for each patient [10]. In another investigation Horsky et. al. found that the absence of multiple system safeguards to check for the type of drug and the dose at successive stages of the medication process contributed to a serious error [11]. In primary health care, which has significantly higher rates of computerization than hospitals, Gandhi et. al. estimated that two out of three prescribing errors could be prevented with safety features that would provide dosage information [12]. These recent studies underline an urgent need to proactively engineer safer systems as the roll out of e-PS accelerates worldwide [13]. Safety is a neglected area as these systems are currently not regulated and there is little formal guidance to inform policy for regulating ICT in healthcare [14-16]. Efforts to investigate the safety of such systems currently rely on an ad hoc combination of methods to retrospectively reconstruct events which led to accidents [6, 8, 11].

In this paper we present a systems approach to studying the complex interactions which give rise to medication errors when e-PS are used in patient care. Developing an accident model, which describes the processes that may lead to failure when an e-PS is used in routine clinical work, should facilitate safer systems design. Such in silico

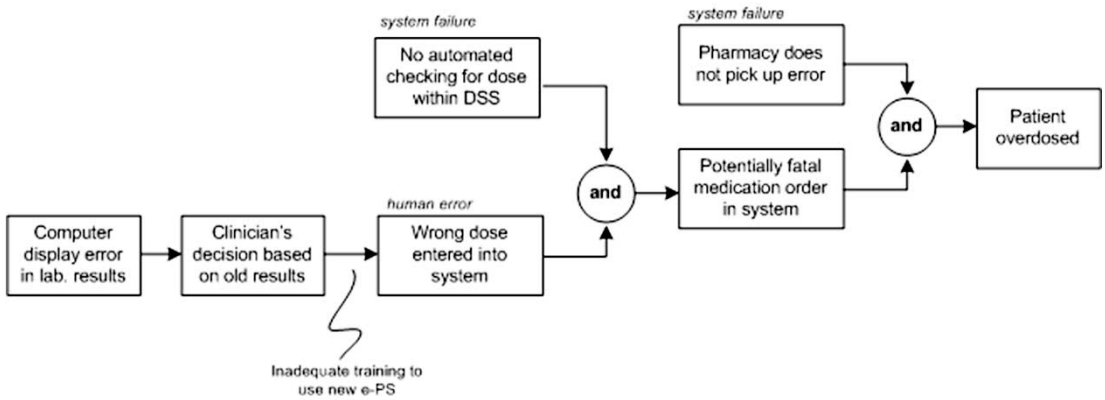


Figure 1 - An accident model for an e-prescribing system (based on case study findings in Horsky et.al. [18])

models provide a safe test bed for exploring complex feedback interactions, and allow system designers to simulate the impact of alternative designs in the clinical environment. Firstly we describe the nature of e-PS failures and review current methodologies to engineering safer systems. We then present an approach to developing an accident model to describe e-PS failures in silico.

The nature of e-PS failures

Efforts to engineer safer e-PS hinge on understanding how accidents occur. Electronic medication management systems support clinicians in complex tasks related to the treatment of disease in individual patients. For example, in the management of chronic disease such as asthma, medications are tailored to individual patients depending on stage of disease, allergies and side effects of drugs. Clinicians' decision-making styles vary within the different specialties and by professional cultures. A significant proportion of the fine-scale decision-making tasks undertaken by clinicians to treat individual patients are complex and cannot be fully standardized [17]. Consequently there is a high potential for human error when using computerized systems, which impose highly standardized human-computer interactions. In the case of medications management, pre-existing *human-human* interactions which impact *human-computer* interactions are particularly significant [18], therefore any accident model must capture task interactions critical to the process. Additionally, the context within which work is undertaken also influences user interactions with e-PS. A number of variables including professional culture and organizational factors give rise to variability in work practices when e-PS are used in a highly socio-technical [19, 20] environment such as healthcare, contributing to system failure. Therefore any effort to analyze medication errors must capture the process, context and task interactions associated with e-PS use.

Approaches to engineering safe systems

An accident model is a representation of the processes which give rise to failures within a system [21] (Figure 1). Such models are used to examine accident causing variables and their associated interactions. Capturing the underlying causes and patterns of failures within an accident model provides a basis upon which to identify hazards and to design interventions to prevent or reduce the impact of accidents. There are no accident models that specifically account for the failure of ICT systems in healthcare. Current efforts to improve safety rely on accident models developed in aviation, defense, nuclear power and the manufacturing sectors, which have pioneered research in systems safety. The Safety Science literature distinguishes three main types of accident models, sequential, epidemiological and systemic accident models [21].

Sequential accident models

The majority of traditional approaches to engineering safer systems are based on sequential models in which accidents are considered as the result of a sequence of events occurring in a specific order. Failure Modes and Effects Analysis is the most popular sequential approach for investigating large technical failures [22]. Although sequential models predominantly focus on technical failures, accidents resulting from human error are considered as any deviation from a standardized operation procedure, much the same way as a discrete technical failure. Probabilistic Risk Assessment [22] which examines the combinations of risk associated technical failures and human errors is being applied to model high-impact, low frequency iatrogenic injury events in medical care e.g. anesthesia patient risk. However, sequential models are inadequate for describing human errors in complex tasks and system failure where direct causal relationships between variables cannot be established e.g. organizational culture and policy.

Epidemiological accident models

Recognizing the difficulty in establishing direct causal relation-ships between events, epidemiological models describe accidents as the outcome of a combination of variables. In this approach accidents occur when a sufficient number of variables, some manifest and some latent, come together in space and time. Reason's Swiss Cheese Model is the best known epidemiological model in which system failures are identified within a causal sequence of events [23]. In com-

parison with sequential approaches, epidemiological approaches such as the Root Cause Analysis [22] consider more than just the proximate events that preceded an accident, therefore they are useful for investigating major accidents but are less useful for smaller adverse events in which system failure is difficult to isolate.

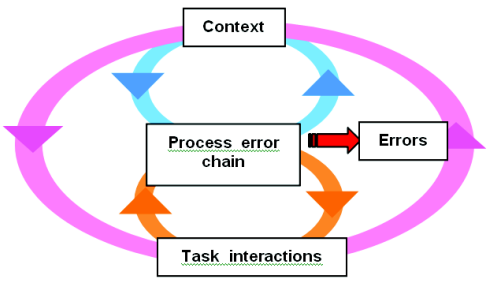


Figure 2 - A multilevel model of the process, context and task interactions which give rise to medications errors

Sequential and epidemiological accident models which examine discrete failure events and focus on causal relationships between variables have limited applicability in engineering safe e-PS. These techniques were designed to assess risks associated with discrete component failures in industrial systems such as a large process plant. In such tightly coupled technical sys-tems, human interaction with machines is broken down into well-defined sequence of events where operational procedures can be standardized. Human error is then considered in terms of deviations from standardized procedures. Healthcare work in contrast is complex, individualized, highly contextual and dynamic [17]. A significant proportion of transactions cannot be completely standardized and therefore human error cannot be fully represented within sequential and epidemiological models.

New systemic approaches to accident modeling

In comparison with sequential and epidemiological accident models, which focus on deterministic causal mechanisms of failure events, systemic models attempt to capture the dynamic behavior of a system. Accidents are viewed as emergent phenomena within complex systems. Hazard analysis in this approach consists of identifying and monitoring the sources of variability that give rise to human error and system failure resulting in accidents. The Functional Resonance Accident Model presented by Hollnagel et. al. is a simple systemic model based on stochastic resonance [21]. In this model, accidents arise from the functional couplings and their inherent variability within complex socio-technical systems. Highly variable hazardous interactions are identified so that barriers can be designed to control their variability and monitor performance. Although this model was specifically developed for highly technical aviation systems in which many of the control actions to maintain safety are automated, the availability of software tools to simulate the model make it a candidate for further evaluation.

Rasmussen's behavioral model is another systemic approach that could be useful for describing the inherent variability in work practices associated with e-PS [25]. This model describes a safety envelope within which a system must operate, defined by four key boundaries to (i) economic failure (ii) unacceptable workload; (iii) operational procedures and (iv) safety boundary beyond which accidents occur. It is essentially an overarching framework to integrate sequential and epidemiological models. Its strengths lie in the fact that it explicitly considers interactions among all levels of the organizational hierarchy to monitor and maintain safety.

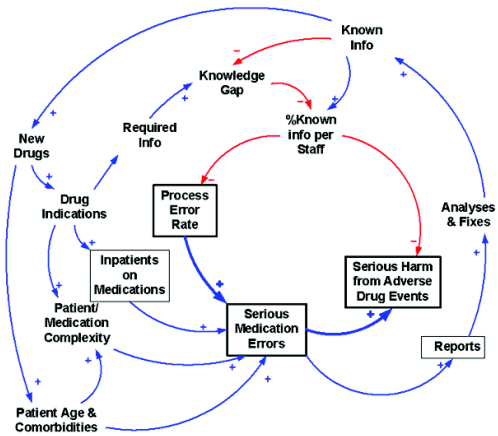


Figure 3 - A causal loop diagram of the process and context interactions associated with medication errors [24]

Perhaps the most sophisticated recent systemic model is that of Leveson which considers safety as a control problem [26]. In this model accidents result from inadequate control and enforcement of safety constraints on a system because (i) hazards are not known, (ii) control action is not adequate or the wrong action is performed, (iii) inconsistencies between process models used by the automation or human (mental models) and the actual process, (iv) missing or inadequate feedback. Similar to Rasmussen's model, this approach views accidents as consequences of socio-technical interactions among all levels of the organizational hierarchy that violate system safety constraints. But it is superior in terms of its formalization based on control theory, a very well developed discipline within Electrical Engineering. When applied to a complex aircraft collision avoidance system this method was found to be more complete than a fault tree analysis [27]. It was also effective in an investigation of a Canadian public water safety system revealing complex couplings between different levels of organization that were responsible for a negative safety culture and ignorance of basic science [27]. Systemic accident models that account for failures in highly automated industrial systems are not directly transferable to healthcare where automation largely supports humans in making complex decisions. However they provide a framework to examine complex relationships at multiple levels of the process, context and task interactions that contribute to e-PS errors.

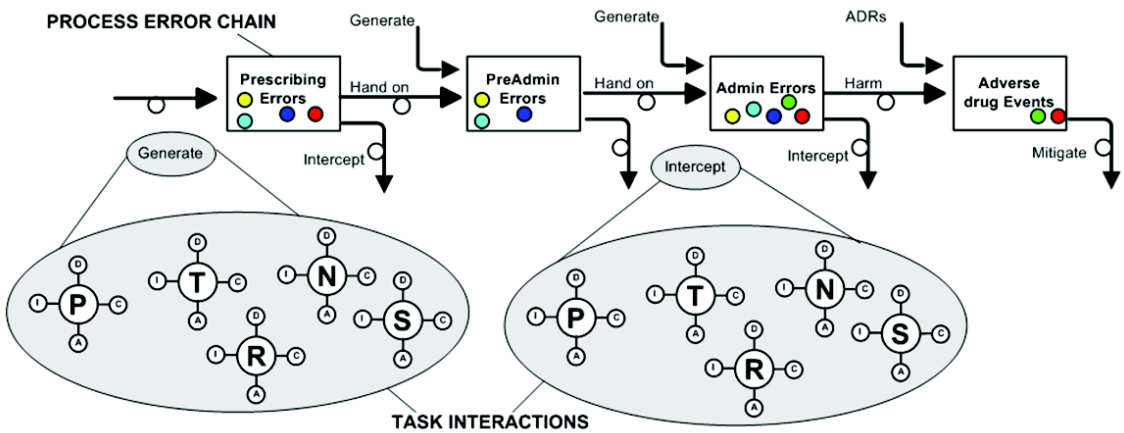


Figure 4 - Task interactions critical to error generation and interception, represented as agents and linked to a process error propagation and mitigation chain. Individual patient type errors can be tracked using patients

A multi-method multilevel accident model for e-PS failures

We propose to use dynamic systems modeling [28] to bring together a comprehensive multilevel model of the process, context and task interaction variables which give rise to human error and system failure when e-PS are used routinely in hospital inpatient care (Figure 2).

Process error chain & context interactions

Based on a systemic approach to accident modeling [26], difference equations and feedback interactions are used at an aggregated level to describe the pattern of interaction between medication management processes and the context that is rele-vant to error generation, interception and transmission. We employ a process chain of error events and harm, based on the framework developed by Bell et. al. [29] for e-PS impacts, and by Anderson et. al. [30] to represent medication error generation, interception and propagation within key prescribing activities e.g. order transmission dispensing, supply and ad-ministration. Bell's framework of e-PS impacts explicitly describes the inputs, outputs, resources and information required at each of the five steps of prescribing, transmission, dispensing, administration and monitoring within the process chain. Anderson's model specifically illustrates the sources of error within this chain. We capture the long-term organizational context that produces errors using a model developed by McDonnell et. al. [24] (Figure 3) which describes the complex interactions of patients and clinicians, information, medications, work practices and infrastructure and policies within the medications management process.

Process & task interactions

Agent based methods can represent task interactions critical to error generation and interception at the disaggregated indi-vidual level. Agents, such as clinicians and patients, are linked to the process error chain (Figure 4). Explicit rules of interaction guide the overall behavior of agents within the model, and errors become an emergent property of the modeled system, rather than being explicitly identified in the model. Key agent actions include informing, deciding, acting and communicating. For example, errors could be generated within prescribing tasks when (i) a specialist is not adequately informed and their decision is based on an old laboratory result, or (ii) an interruptive environment imposes cognitive loads resulting in a wrong decision, or (iii) from ambiguous computer screens that do not allow residents to enter orders correctly (action), or (iv) medications changes are not communicated in a timely manner to nurses. Individual patient errors are tracked by including patients within the model. The approach serves to identify error modes due to misperceptions of the situational context (informing role), cognitive errors (deciding role), task execution (acting role) and communication especially during transmission to downstream processes.

We will undertake a process to systematically assess validity of the model (a) conceptual validity using expert opinion, research evidence and study data (b) structural validation through formal qualitative inspections to confirm structure, parameters, extreme conditions and dimensional consistency. The model will also be examined quantitatively under extreme conditions to assess behaviour sensitivity, boundary adequacy and phase relationship (c) behavioural validity using behaviour pattern tests (d) simulation verification and (e) replication. The validation and verification of the simulation model will be based on Sterman's evaluation framework [28]. The model will be calibrated using data from the controlled experiments and observational studies.

Discussion & conclusion

There are no empirical accident models to provide a basis for engineering safe ICT systems for complex socio-technical systems such as healthcare. Sequential, epidemiological and systemic models developed within

defense, aviation, nuclear power and industrial applications are not directly transferable to healthcare, which is characterized by complex interactions that are not easily standardized. We have presented a framework for a systems approach to developing a novel accident model to describe e-PS accidents in healthcare. By bringing together a multi-level multi-method model of process, context and task interaction variables we will develop an accident model that accounts for human error and system failure mechanisms associated with complex fine-scale tasks unique to decision-making in dynamic environments such as healthcare. The novelty of this approach lies in combining the strengths of multiple systems and agent-based modeling methods within a simulation environment to examine complex interactions between multiple parameters otherwise not possible with traditional techniques. The next stages of this work will involve identification of variables to calibrate the model. We will then experiment with safety interventions to minimize the impact of human error and system failure on medication errors.

Acknowledgments

FM is supported by the Faculty of Medicine, University of New South Wales; JW is supported by an Australian National Health and Medical Research Council (NHMRC) CD grant.

References

[1] US Institute of Medicine: Crossing the quality chasm: A new health system for the 21st Century. USA: National Academy Sciences Press; 2000.
[2] National Health Service: Reports by the chief pharmaceutical officer. Building a safer NHS: Improving safety medication. London: National Health Service.
[3] Leap Frog Group. Computer assisted order entry fact sheet. Washington: Leap Frog Group; 2001.
[4] Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2004;11(2):104-12.
[5] EFMI. Bad Health Informatics Can Kill: Working Group for Assessment of Health Information Systems of the European Federation for Medical Informatics(EFMI): iig.umit.at/efmi/ [14 Nov 2006].
[6] Han YY, Carcillo JA, Venkataraman ST, Clark RS, Watson RS, Nguyen TC, Bayir H, Orr RA. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. Pediatrics 2005;116(6):1506-12.
[7] Gesteland PH, Nebeker JR, Gardner RM. These are the technologies that try men's souls: common-sense health information technology. Pediatrics 2006;117(1):216-7.
[8] Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, Strom BL. Role of computerized physician order entry systems in facilitating medication errors. JAMA 2005;293(10):1197-203.
[9] US Pharmacopeia: MEDMARX 5th Anniversary Data Report: A Chartbook of 2003 Findings and Trends 1999ñ2003; 2004.
[10] Nebeker JR, Hoffman JM, Weir CR, Bennett CL, Hurdle JF. High rates of adverse drug events in a highly computerized hospital. Arch Intern Med 2005;165(10):1111-6.
[11] Horsky J, Kuperman GJ, Patel VL. Comprehensive analysis of a medication dosing error related to CPOE. J Am Med Inform Assoc 2005;12(4):377-82.
[12] Gandhi TK, Weingart SN, Seger AC, Borus J, Burdick E, Poon EG, Leape LL, Bates DW. Outpatient prescribing errors and the impact of computerized prescribing. J Gen Intern Med 2005;20(9):837-41.
[13] Fox J, Das S. Safe and sound: artificial intelligence in hazardous applications /. California: AAAI Press/MIT Press; 2000.
[14] Coiera EW, Westbrook JI. Should clinical software be regulated? Med J Aust 2006;184(12):601-2.
[15] Australian Government Department of Health and Ageing Therapeutic Drugs Administration: www.tga.gov.au/devices/devices.htm. [cited Oct 24]
[16] US Food and Drug Administration: Regulatory Requirements for Medical Devices: www.fda.gov/cdrh/manual/510kprt1.html. [cited Oct 24];
[17] Bar-Yam Y. Making things work: solving complex problems in a complex world. Newton: NECSI-Knowledge Press; 2004.
[18] Coiera E. Interaction design theory. Int J Med Inform 2003;69(2-3):205-22.
[19] Berg M. Patient care information systems and health care work: a sociotechnical approach. Int J Med Inform 1999;55(2):87-101.
[20] Coiera E. Four rules for the reinvention of health care. BMJ 2004;328(7449):1197-9.
[21] Hollnagel E. Barriers and accident prevention. Aldershot: Ashgate; 2004.
[22] Dhillon BS. Human reliability and error in medical system. Singapore: World Scintific Publishing; 2003.
[23] Reason J. Human Error. Cambridge: Cambridge University Press; 1990.
[24] McDonnell G, Heffernan M. The Dynamics of Hospital Medication Errors: A Systems Simulator Testbed for Patient Safety Interventions. In: 23rd International Conference of The System Dynamics Society July 17 - 21, Boston.; 2005.
[25] Rasmussen J. Risk management in a dynamic society: a modelling problem. Safety Science 1997;27(2/3):183.
[26] Leveson N. A new accident model for engineering safer systems. Safety Science 2004;42(4):237-270.
[27] Leveson N. Model-Based Analysis of Socio-Technical Risk: Technical Report, Engineering Systems Division, Massachusetts Institute of Technology.; 2002 June.
[28] Sterman JD. Business Dynamics: Systems Thinking and Modeling for a Complex World. Irwin: McGraw-Hill; 2000.
[29] Bell DS, Cretin S, Marken RS, Landman AB. A conceptual framework for evaluating outpatient electronic prescribing systems based on their functional capabilities. J Am Med Inform Assoc 2004;11(1):60-70.
[30] Anderson JG, Jay SJ, Anderson M, Hunt TJ. Evaluating the capability of information technology to prevent adverse drug events: a computer simulation approach. J Am Med Inform Assoc 2002;9(5):479-90.

Address for correspondence

Centre for Health Informatics,
University of New South Wales,
Sydney 2052, Australia.
Email: f.magrabi@unsw.edu.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Securing Chemotherapies: Fabrication, Prescription, Administration and Complete Traceability

Stéphane Spahni[a], Christian Lovis[a], Monique Ackermann[b], Nicolas Mach[c], Pascal Bonnabry[b], Antoine Geissbuhler[a]

Service of Medical Informatics, Department of Radiology and Medical Informatics[a] Pharmacy[b,] Service of Oncology, Department of Internal Medicine[c] University Hospitals of Geneva, Switzerland

Abstract

Decision support, order entry, drug and care administration with their respective documentation cannot be seen as independent actions, especially in term of medical approach and patient safety. Chemotherapy treatment is a good example to illustrate the various implication of technology information in these multifaceted and intricate processes. Chemotherapy administration can be a highly complex process. It can take place over a variable period of time, from hours to months. Chemotherapies can be produced specifically for a given patient and can have dramatic effects in case of errors. Chemotherapy treatment will depend from various information including patient specific data such as temperature, weight or laboratory or drug specific knowledge such as side effects or administration directives. At the University Hospitals of Geneva (HUG), processes reengineering accompanied with new applications covering the whole chain of the processes involved in chemotherapy treatment (prescription preparation, administration, control) have been developed. This paper presents the overall approach leading the computerization of these processes and the first evaluations about the potential benefits of the computer-aided controls during the administration phase.

Keywords:

hospital information systems (HIS), GS1, RFID, pharmacy, chemotherapy.

Introduction

The need for increased safety and efficiency in care production is an important goal in healthcare [1]. Optimization of care production and a better efficiency of care logistics while improving quality and safety is however a challenging goal. Information technologies allow to meet both goals, increasing both efficiency and safety [2, 3]. However, reaching these goals implies to embed information technologies at all steps of the process, to identify clearly all actors and objects involved and, usually, to reengineer deeply these processes, the associated procedures and the way the are (often historically) done. To illustrate this, we present the developments made around chemotherapy treatments at the HUG.

The HUG is a consortium of hospitals in four campuses and more than 30 ambulatory facilities in the state, comprising more than 2,000 beds, 5,000 care providers, over 45,000 admissions and 750,000 outpatients' visits each year. It covers the whole range of in- and outpatient care, from primary to tertiary facilities. The HUG is the major public healthcare facility in the Geneva region and the near France. The in-house developed computerized patient record (CPR) is used in all facilities and runs on more than 4,500 PC's. Over than 20,000 records are open every day by more than 4,000 care providers from all functions [4].

Chemotherapeutic agents are prescribed for various oncological and haematological disease states. Although they are considered to be the treatment of choice for many cancers, these medications are associated with serious and potentially life-threatening side effects. The toxicities of these anti-cancer drugs, the multidisciplinary actors involved in the whole treatment process, create a very high risk of devastating medical errors. In an effort to minimize the potential for chemotherapy-related errors, the HUG have spent the last years performing important developments and prospective risk analysis to lower the criticality of the whole process [5]. As part of the results, the preparation of all chemotherapies administered at any of the facilities of the HUG has been centralized at the pharmacy. The centralization guarantees the processes, the quality and the traceability of the preparations and their components and, very important, the safety of the operators. This first phase allowed justifying the need and the pertinence of a unique description of substances used and their associated protocols. One major task was the creation of a global database containing all substances, materials and description of procedures that can be used for in-house preparations. This could not be done without reaching consensus between all actors involved. Building the protocols is one of the very complex tasks that have been made in close collaboration with and under the supervision of specialists. This database can then be used by the applications developed for requesting new preparations, managing the pending requests, organizing the concrete preparations and managing the traceability before, during and after the production process.

The prescription side

At the prescription side, we developed a family of tools including components to help the creation and management of specific protocols, to use them for given patients, but also to help the follow-up of all patients getting chemotherapies. With the correct right accesses, users can see the list of patients and their respective chemotherapy protocols. For each patient, one can see various information, such as the protocol type, the number of cycles involved, the first day and the next day of chemotherapy. These features help oncologists to follow their patients, the current status of all running protocols, but it is also a powerful tool for evaluation of the treatment, the prevention of side effects and the efficiency of this prevention.

Figure 1 - chemotherapy protocol for a given patient

At the time of prescription, physicians can choose the appropriate protocol and use it with the data pertaining to the correct patient. The system will help adjusting doses and chronology of procedures, as well as all elements pertaining to the selected chemotherapy regimen, including the fabrication of the chemotherapy if required (see Figure 1).

The system can produce alerts in some situations, for example when overlapping dates and regimens. Other alerts might be added in the future, such as warning if the renal function is worsening, or the white blood cells counts too low.

The pharmacy side

In many cases, the regimen associated with a protocol does not exist per se commercially, but must be produced specifically by the HUG's central pharmacy.

Figure 2 - the main page of the pharmacy system for drug production

The tools developed for the pharmacy allow the management of most of the logistics needed to produce drugs from raw substances. This includes reception of the raw substances and their identification/validation, tagging, stocking and localization, lot follow-up and usage. They also ensure a complete traceability of all actions on these substances or their derivates up to the final product preparation and its distribution. The production of products must follow strict production protocols - distinct from prescription protocols – that includes materials, procedures, validation steps and safety behaviours. This production includes specific chemotherapy treatments, but also a wide range of in-house made products, such as anti-cough syrup or disinfectants, amongst others. The computerized production protocols can be adapted according numerous parameters, ranging from patient specific data to mass production variables. Drugs validation and laboratory analyses are often made both at the level of raw substances and on the final product (Figure 2). These analyses ensure the quality of the raw components and of the final product used for patient care. As long as a product, raw or final, has not been validated, it is kept in "quarantine" and cannot be used. The system allows following of all lots, production and expiration dates, suppliers, end users, remaining stocks, and the position of the product in the production workflow (Figure 3). It ensures standards in the production chain and increases operators' safety. In parallel, a computerized high precision balance aiming at reducing the risk of errors during the fabrication is used by a specific tool for supporting the fabrication of specific preparations (Ca[1]to®). This tool automatically gets the various volumes of the preparations to be realized. These

data are used to ease the validation of each step by controlling the plausibility of the result.

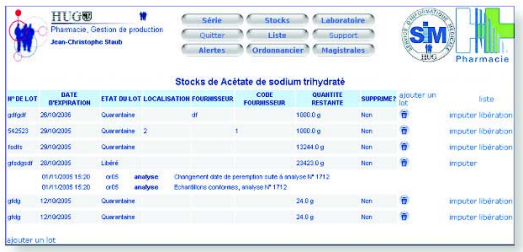


Figure 3 - Example of workflow and stock management

Drug administration – the nursing side

One of the complex tasks of computerization of this process involves the administration part. Before the chemotherapy is administrated, complete instruction will be printed in the ward for nurses, including information about drugs, sides-effects and their prevention such as nausea, and precautions for the administration itself. These precautions will cover many aspects, such as solutions incompatibilities, administration temperature, administration route or protection against light (Figure 4).

The last step in the process is the bedside traceability of the administration of the drug to the patient. In order to be able to properly trace all steps of administration, it has been necessary to tag all infusions and actors, the nurses involved and the patient. This phase is currently ready for its initial deployment in a pilot ward, but is delayed due to logistics problems with the labels. A solution based on Pocket PC is being realized. The goals of this module are the following:

- Verification that it is the right preparation for the right patient;
- Automatic control that the preparation is still administrable (used-by date);
- Verification at the bedside that no stopping condition have been issued meanwhile;
- Complete traceability up to the "last mile".

These functions require the correct identification of the three "partners" involved: the patient, the nurse and the preparation. While there are many ways to identify the partners, proprietary or non proprietary, we decided to start using international numbering of objects and selected GS1's coding schema (GS1 is the result of the merge of the two organizations EAN – European Article Numbers and UCC – Uniform Code Council). All three partners have therefore their own code that can be read by the Pocket PC and validated on-line according to the information stored in the Hospital Information System (HIS): the patient and the nurse have their own permanent GS1 codes and each preparation produced by the pharmacy receives an unique identification code enabling the tracking in the institution. All GS1 codes are using the GS1-128 encoding scheme.

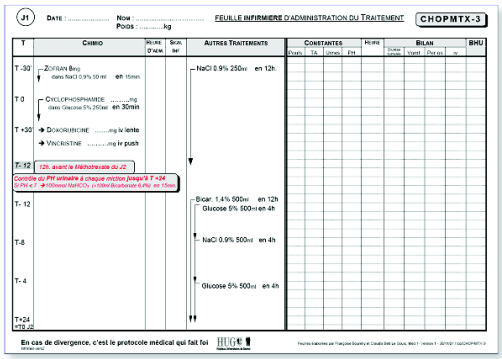


Figure 4 - Administration directives

The GS1 encoding can be used on various transporters, such as linear barcodes, data matrix, radio-frequency tags, etc. The problem with the linear barcodes comes from the size of the printed code: 1 dimensional formats are rather large, and data matrix are not commonly read by classical readers. The labels on infusions and other preparations are devoted to human-readable information, therefore reinforcing the ability for human to cross-check information (Figure 5). For this reason, we decided not to use printable codes. To achieve this, the use of labels with integrated RFID chips has been chosen for the electronic identification of the preparations as well as the patients (using a wristband with a RFID sticker) and the nurses.

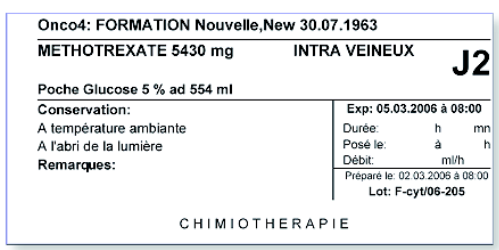


Figure 5 - Self adhesive label with a radio-frequency tag (RFID)

The labels for preparations are printed at the pharmacy before production and joined to the raw material required for the fabrication, while the labels for patients will be produced at the admission desks (they are currently produced at the ward until the trial is complete). This solution enables the use of a unique reader at the bedside for safely getting the identities of the patient, the nurse and the preparation to be administered. The RFID reader is used with a PDA (HP iPaq) with wireless LAN connection to the CPR. Therefore, validation check can be made in real-time and at the bedside. The same device is also used after the start of the administration for e.g. controlling the current administration status after the staff rotation, for indicating

1 Computer Aided Therapy for Oncology (cato). http://www.cato.at

a possible suspension of the treatment and finally for registering the end of the administration.

Figure 6 - The overall administration process at the bedside

Initial evaluation

Evaluation of the impact of the introduction of such a project is important in many respects – benefits in patient safety, acceptance by the users, impact on the procedures and workload, etc. Recent work of Bates [6] shown that the origin of errors in drug prescription and administration are distributed as follows: prescription 49%, re-transcription 11%, preparation 14% and administration 26%. For patient safety, considering that if about 50% of the errors during the prescription, preparation and dispensing phase are caught only 2% of the administering and documenting errors are detected [5, 7], having a reliable evaluation of the impact is crucial. A study has therefore been conducted in order to evaluate the safety benefits and the acceptance of the use of PDAs during the administration phase [8, 9], addressing thus specifically the 26% of drug dispensation errors. This study has yet only been realized "off-site" with 62 users (specialized and non-specialized nurses as well as assistants at the pharmacy). The study was oriented in order to measure the following aspects:

- Benefits of the computer-assisted controls against a fully manual control with or without a paper-based check-list (7 checkpoints: identity of the patient, name of the substance to administer, dose, administration path, date and day of administration use-by date taking into account administration duration, conservation);
- Acceptance of the system by the future users;
- Getting a feed-back regarding the ergonomics of the application.

In order to evaluate the respective benefits of the two helps (checklist, PDA), participants had to check two sets of protocols containing randomized errors (2% and 16% of errors respectively). The results were the following [9]:

Non-specialized persons:

	No assistance	Checklist	PDA
Detection [%]	83.5 %	97.5 %	100 %
Trust interval	78.9 – 87.4	95.1 – 89.9	
Test duration [min]	32	18	8

Specialized persons:

	No assistance	Checklist	PDA
Detection [%]	90.1 %	100 %	100 %
Trust interval	85.7 – 93.6	87.2 – 100	
Test duration [min]*	32	18	8

* No significant differences in the mean test duration were noted between the two categories of persons.

As one can see, the results of the study are very promising for the future use of the new application:

- While the use of the check-list already significantly reduces the number of undetected errors, it is only with the PDA that non-specialized staff detected 100% of the problems;
- The time spent for doing the controls is reduced by a factor of up to 4 when using the PDA.
- Between 73% and 80% of the persons preferred to use the PDA in various situations while they are only 13% to 20% preferring to use the check-list.

We are therefore very confident that the new system will bring a valuable contribution to patient's safety and daily work of nurses. Further evaluations will of course be performed after the whole system is in production.

Discussion and conclusion

Implementing such a system is a challenging goal, merging virtual and real worlds. It involves numerous actors and cannot be achieved without managing carefully organisational impact. Analyzing the complete process of chemotherapy treatment, from the design of a protocol, to the prescription, the production, the administration and the follow-up as a coherent and shared workflow is an important paradigm in improving both safety and efficiency of the process.

On the clinician side, this process has initiated a global formalization and sharing of protocols and guidelines. Except special cases, the prescription of drugs regimens not agreed within protocols is not more possible. Only adjustments pertaining to patient characteristics are allowed; doses are then validated according to several variables that cannot be over

passed. The prescription has gained in transparency and readability. In the pharmacy, benefits are numerous, from better management of raw substances and traceability to increased safety and standardization of operators work. For nurses, validation of the administration and confidence in the overall process should be increased when the final phase will be introduced. By now, there is already a clear benefit due to the standardization, the completeness and the readability of treatment directives.

The system has been well accepted by all actors involved. The importance of the quality of the information is well recognized, and the tools do offer a significant improvement. However, it must be emphasized that the formalization and the validation of all processes, including each protocol, require a significant amount of time, especially from oncologists. But we now have reached the state where it is the oncologists who do not have the new applications who are strongly pushing for getting it (at the cost of formalizing and standardizing their protocols).

The initial project for the support of the prescription and production of chemotherapies has been successfully deployed and is well accepted by its users. The second phase, the bedside validation of the administration of the medications to the patients is in development and should be used in a pilot ward within a few months. We expect to have a significant improvement to the overall security of the care of the patients thanks to several achievements: better documentation in the patient record, suppression of the hand-written orders, controls at the prescription, suppression of multiple retranscription, production and administration levels, etc. The use of Pocket PCs and radio-frequency technologies at bedside is expected to grow as real-time application and complete traceability are being progressively installed.

The patient will be the final beneficiary of the system, getting a globally improved process with increased safety and traceability. Further studies will be held once the bedside administration validation system will be in production.

References

[1] ITAC. Report to the President, Revolutionizing Health Care through Information technology. President's Information Technology advisory Committee June 2004.

[2] Bates DW, Cohen M, Leape LL, Overhage JM, Shabot MM, Sheridan T. Reducing the fre-quency of errors in medicine using information technology. J Am Med Inform Assoc 2001;8(4):299-308.

[3] Ruland CM. Improving patient safety through in-formatics tools for shared decision making and risk communication. Int J Med Inform 2004;73(7-8):551-7.

[4] Geissbuhler A, Lovis C, Lamb A, Sphani S. Ex-perience with an XML/HTTP-based federative approach to develop a hospital-wide clinical in-formation system. Medinfo 2001;10(Pt 1):735-9.

[5] Bonnabry P, Cingria L, Ackermann M, Sadeghi-pour F, Bigler L, Mach N. Use of a prospective risk analysis method to improve the safety of the cancer chemotherapy process. Int J Qual Health Care 2006;18(1):9-16.

[6] Bates DW, Cullen DJ, Laird N et al. Incidence of adverse drug events and potential adverse drug events. JAMA 1995; 274: 29-34.

[7] Bates DW. Using information technology to re-duce rates of medication errors in hospitals. BMJ 2000; 320: 788-91.

[8] Bates DW, Nathan S, Cullen DJ et al. The costs of adverse drug events in hospitalized patients. JAMA 1997; 277: 307-11.

[9] Baalbaki R. Mesure de la fiabilité du contrôle ul-time avant l'administration des médicaments aux patients. Diploma work presented at the Science Faculty of the University of Geneva, Switzer-land, July 2006.

Address for correspondence

Stéphane Spahni
Medical Informatics Service
Radiology and Medical Informatics Department
University Hospitals of Geneva
24 rue Micheli-du-Crest
CH-1211 Geneva 14
Email: Stephane.Spahni@hcuge.ch

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Multitasking by Clinicians in the Context of CPOE and CIS Use

[a]Sarah Collins, [a,b,c] Leanne Currie, [c]Vimla Patel,
[a,c]Suzanne Bakken, [c]James J. Cimino

[a] *Columbia University School of Nursing, New York, New York, USA*
[b] *New York Presbyterian Hospital, New York, New York, USA*
[c] *Department of Biomedical Informatics, Columbia University, New York, New York, USA*

Abstract

Interest in studying distractions and interruptions in the context of clinician workflow has increased in light of evidence that these events may negatively impact patient safety. Additionally, many recent informatics-based studies analyze computer provider order entry (CPOE) and clinical information system (CIS) implementation and its effects on clinician workflow. This study expands the development and use of a taxonomy to characterize distractions to clinicians and their subsequent actions in the context of CPOE/CIS use. We found a total of 75 distracting events in 406 minutes of observational data of which 32 led to interruptions and 30 led to continued multitasking. The above primary actions led to 5 tasks not completed and 4 episodes of clinician's lack of recall. Differences in the distribution of the source of distractions and primary action of the distracted clinicians may be a function of clinical setting and clinician type using the CPOE/CIS. Nine secondary actions, potentially resulting in a slip or a mistake, suggest that CPOE may necessitate different forms of safety nets than traditional clinician communication.

Keywords:

communication, distraction, interruption, computer provider order entry, clinical information system

Introduction

Multitasking is a valued skill in the clinical setting allowing for the efficient execution of daily activities, yet, may fail to be an effective mechanism in clinical practice when it leads to cognitive overload [1]. Clark and Brennan [2] noted that the properties of a communication medium impose constraints on the communication process; therefore, the design of communication modalities should account for these constraints [3].

During coordinated activities, such as communication amongst the health care team, responsible clinicians must establish common ground. Common ground is defined as similar experiences, beliefs, and knowledge and is necessary to ensure that clinicians' mental models reflect each others' needs within the context of the task and current situation [4]. Coordinated communication may serve as a "rescue" to account for the fact that the introduction of each new interaction can detract from the clinician's finite cognitive resources [4]. When cognitive resources are exhausted, the average amount of attentional resources, also known as working memory, available to any single interaction may be reduced [5].

Computerized Provider Order Entry (CPOE) systems attempt to alleviate clinician cognitive overload through providing an organized electronic format for work activities. CPOE systems can have several types of decision support functions such as automated alerts and reminders; however, excessive alerts have been shown to cause cognitive overload and alert fatigue [6]. Alert fatigue occurs when the number of reminders and alerts are perceived too be excessive. This number often varies from clinician to clinician and may cause clinicians to override both critical and non-critical alerts, compromising the desired safety effect of integrating decision support into CPOE [6].

The benefits of CPOE implementation likely outweigh the unintended consequences [7]. However, due to the asynchronous channel's inability to form a mental model of the cognitive needs of the clinician, new ways to detect distractions, interruptions and multitasking in the setting of CPOE and general clinical information system (CIS) usage are needed. Furthermore, other clinicians within the same workspace may be equally vulnerable to similarly reduced attentional resources. This study describes the use of a taxonomy to characterize and analyze distractions and subsequent actions in the setting of CPOE/CIS usage.

Background

Studies characterizing interruptions and distractions during clinician workflow continue to demonstrate their prevalence and significance to the health care work environment's culture [8,9]. Interruption rates consistently approach 30% of all clinical communication, in many settings. On the other hand, interruptions serve as the most frequently used method of communication in the health care environment and are thus considered a beneficial activity [8]. Of great concern however, up to 43% of medication errors have been attributed to distractions such as interruptions [10]. CPOE systems show great promise in their potential to increase patient safety and have been

Selected for best paper award.

shown to reduce medication errors up to 81% [11]. However, unexpected "silent errors", i.e., latent errors resulting from mismatches between clinician workflow and CPOE or CIS development and deployment, have emerged as potentially contributing to medical errors [7].

Table 1 – Taxonomy of distractions during CPOE use

Distraction with multitasking (Event): Period when a distraction causes a subject to interact in two or more concurrent communication events [9].

Primary Actions	Definition	Examples				
		User	**Source**	**Original Task**	**Distraction Quote or Episode**	Secondary **Action**
Interruption (I)	Cessation of productive activity before the current task was completed for an externally imposed reason [13].	MD	*C*	Ordering urine electrolytes	Difficulty navigating CPOE order process; sought assistance from second clinician but was unsuccessful	*IT*: User stopped after 1 minute and 5 seconds of unsuccessful attempts.
Continued Multitasking (CM)	Continued interaction in two or more concurrent communication events or tasks.	MD	*SC*	Ordering Mucomyst medication	"Is that the patient's chart you have?"	*LR*: Difficulty recalling Mucomyst dose previously stated by another clinician.
Deferred Task (DT)	Acknowledged distraction stimulus from external source that was not followed by cessation of original task or continued multitasking.	MD	*SC*	Ordering patient laboratory draws	"Did patient X ever get the CT and biopsy?" (different patient than active patient record on CPOE screen)	*IT*: User acknowledged distraction with indication to follow-up, yet follow-up was not observed.

Secondary Actions	**Definition**
Lack of Recall (LR)	Inability to quickly recall previously verbalized information about the task (could result in slip or mistake [16]).
Incomplete Task (IT)	Either original task and/or interruption task not completed during the observed session (could result in slip or mistake [16]).
Change in Plan of Care (CPC)	Information of a new plan interrupts a task and causes the interrupted task to be irrelevant and discarded (could result in slip or mistake [16]).
Source of Distraction Event	**Definition**
Information Need (IN)	Required data to answer a question is not known.
Synchronous Communication (SC)	Two parties exchange messages across a communication channel at the same time (e.g. face-to-face, telephone) [17].
Computer (C)	Distraction caused by technical problems (e.g. frozen user interface) or CPOE/CIS usability difficulties.

In an earlier study, we developed a taxonomy to characterize distractions and interruptions during the use of CPOE in a medical intensive care unit (MICU) [12]. The taxonomy extended the work of Coiera et al.'s definition of distraction with multitasking [9] and Flynn et al.'s definition of an interruption [13]. Our previous study found that information needs accounted for 55% of the distraction events detected and clinical communication accounted for 40% [12]. Distraction events may be attributed to more than one source. For example, a clinical communication may be related to an information need. Additionally, 14% of distractions were attributed to computer problems. For example, while the observed MICU resident was changing a CPOE oxygen order, another clinician distracted him with an information need. In this instance, the second clinician asked for information that the first clinician needed to obtain from the CIS. The distraction interrupted the resident, the resident answered the question, and then lack of recall of the original intended oxygen order occurred for 1 minute and 6 seconds [12].

In a second study, as part of the Infobuttons project at Columbia University, 51% of clinician information needs went unmet [14]. Consistent with Coiera et al.'s findings that clinicians use synchronous channels of communication more frequently, Currie and colleagues found that for

76% of information needs concerning domain-related questions, the individual in need of the information utilized a person to answer the question (rather than a paper or computer-based resource) [14]. The health care setting, rich with information needs and with an apparent preference for face-to-face communication sets up an interrupt-driven environment [15] which has previously been shown to both compromise patient safety [10] and to provide an opportunity for coordinated activity "rescues" [4]. However, to date, these phenomena have been studied in the absence of CPOE/CIS, therefore this study characterizes and analyzes interruptions and distractions to clinicians in the context of CPOE/CIS use.

Methods

Morae ™ software served as the portable usability lab to collect all observational data during randomly selected periods of time [14]. We used a Taxonomy of Distractions During CPOE use which was developed for earlier work on this topic to characterize distraction and interruption events during observational data of medical resident rounds in the MICU (See Table 1) [12]. We developed the taxonomy by iteratively using deductive and inductive methods to characterize our observational data. This was congruent with the hybrid method to categorize interruptions and activities as described by Brixey et al. in the HyMCIA study [18]. For this study, we used the taxonomy and extended our sample to residents and other clinician's who were using CPOE/CIS on inpatient medical/surgical floors collected for the Infobuttons project at Columbia University Medical Center (CUMC).

Table 2 – Time and events by type of clinician for combined rounds and non-rounds data

Clinician	Minutes (%)	#Distraction	Events/Hr	Interruption (%)	Continued	Multitask-Deferred Task (%)
MD MICU Rounds	93 (23)	22 (29)	8	10 (31)	3 (10)	9 (69)
MD Non-rounds	158 (39)	41 (55)	28	13 (41)	25 (83)	3 (23)
RN Non-Rounds	94 (23)	8 (11)	5	5 (16)	2 (7)	1 (8)
PT/OT Non-Rounds	45 (11)	2 (2.5)	2.5	2 (6)	0	0
Student Non-Rounds	16 (4)	2 (2.5)	7.5	2 (6)	0	0
Total	406	75	Avg 10	32	30	13

The coding schema, as shown in Table 1, which was used to characterize distraction events, includes the initial event of Distraction with Multitasking when the clinician engages in another interaction in addition to the current use of CPOE. The initial event is followed by a primary action of the distracted clinician: an Interruption, a Deferred Task, or Continued Multitasking. The primary action may or may not be followed by a secondary action: Lack of Recall, Incomplete Task, or Change in Plan of Care.

Results

We observed 38 clinicians from a combined data set of MICU rounds and CPOE/CIS use on a medical/surgical unit at the Columbia-Presbyterian campus of New York Presbyterian Hospital. A total of 75 Distraction with Multitasking events were detected in 406 minutes of observational data. Table 2 shows the breakdown of total minutes and events per type of clinician observed.

Distractions during MICU rounds

During 93 minutes of MICU rounds, observational data of a medical resident using CPOE found that a total of 22 distractions occurred, with one distraction occurring on average every 4.2 minutes. Ten of the events were categorized as Interruptions, three as Continued Multitasking, and nine as a Deferred Task. An Interruption preceded one Incomplete Task and two Lack of Recall episodes and one Change in Plan of Care. One Deferred Task preceded one Incomplete Task. Three sources of distraction were identified. It is important to note that these sources are not mutually exclusive; therefore, the sum of the percentages does not equal 100 percent. These sources were as follows: information need events [14] accounted for 12 out of 22 events (55%), clinician communication accounted for 9 out of 22 events (40%), and frozen CPOE user interface screens accounted for 3 out of 22 events (14%).

Distraction during MedSurg non-rounds

The second observational data set consisted of clinicians (physicians, nurses, physical therapists, occupational therapists, and medical students) while using a CPOE/CIS system. The observational data ranged in length of time from one minute to 37 minutes. The recording time represents time when the user was actively using the CPOE/CIS system.

The analysis of 313 minutes of data that was collected identified a total of 53 Distraction with Multitasking events. When examining overall data, on average a distraction event occurred every six minutes. Of these 53 events: 27 events were characterized as a Continued Multitasking primary action with two secondary actions of Lack of Recall; 22 events were characterized as an Interruption primary action with two secondary actions of Incomplete Tasks; four events were characterized as a Deferred Task primary action with one secondary action of Incomplete Task.

Distractions by user type

Sixteen physicians were observed using CPOE/CIS for a total of 158 minutes with 41 identified Distraction with Multitasking events. Of the 41 events, 25 (60%) resulted in a Continued Multitasking primary action and two of those primary actions led to a Lack of Recall secondary action. Thirteen of the 41 events, or 31%, produced an Interruption primary action with one of those actions leading to an Incomplete Task secondary action. Of the remaining three

Deferred Task primary actions, only one led to an Incomplete Task secondary action. An event occurred every 3.8 minutes while a physician was using CPOE/CIS.

Observational data of 13 nurses using CPOE/CIS totaled 94 minutes and identified 8 Distraction with Multitasking events. 62%, or 5 of the 8 events, were followed by an Interruption primary action, with one of those actions leading to an Incomplete Task secondary event. In contrast, 2 (25%) of the 8 events led to Continued Multitasking primary actions. The remaining event was identified as a Deferred Task primary action. An event occurred every 11.75 minutes while a nurse was using CPOE/CIS.

Four physical therapists were observed using CPOE/CIS for a total of 31 minutes and encountered one Distraction with Multitasking event. Observational data collected on one occupational therapist using CPOE/CIS for a total of 14 minutes identified one Distraction with Multitasking event. The two events that occurred during physical therapist and occupational therapist CPOE/CIS use both resulted in an Interruption primary action with no occurrence of a secondary action. On average a physical therapist or occupational therapist encountered an event every 22.5 minutes.

Three medical students were observed using CPOE/CIS for 16 minutes with two Distraction with Multitasking events identified. Each of the two events encountered by medical students was followed by an Interruption primary action and no secondary action. The medical students, on average, experienced an event every eight minutes.

Of the 75 events in the combined data set, 10 resulted in a secondary action (See Table 3).

Table 3 – Total count of primary and secondary actions

Primary Actions		Secondary Actions		
		Lack of Recall	Incomplete Task	Change in Plan of Care
Interruption	N = 32	2	3	1
Deferred Task	N = 13	0	2	N/A
Continued Multitasking	N = 30	2	0	N/A
Total	N =75	4	5	1

An information need [14] accounted for 21 of the 53 non-rounds events (40%). In these cases, 13 of the 22 information needs caused an Interruption, two of which the information need failed to be met and an Incomplete Task resulted. Six of the 22 information needs caused Continued Multitasking, two of which caused Lack of Recall episodes. Two Deferred Tasks resulted from an information need, one of which the information need was deferred and resulted in an Incomplete Task. Of note, 17 of the 21 events resulting from information needs utilized a human resource [14].

Discussion

We observed 6 hours and 46 minutes of clinicians using a CPOE/CIS system in the MICU and in MedSurg. Distractions occurred 75 times during this time period. The MICU rounds data resulted in a greater proportion of Deferred Tasks while the medical/surgical unit non-rounds data resulted in a greater proportion of Continued Multitasking. However, Interruptions as a primary action from a distraction event were found to have high proportions in both the rounds and the non-rounds data. Possibly the nature of the structure, pace and coordinated activity [4] of MICU rounds leads the resident using CPOE/CIS to be less likely to engage in Continued Multitasking. The less structured work of non-rounds CPOE/CIS use may increase the likelihood of engaging in conversation with a colleague while using the CPOE/CIS system, a form of Continued Multitasking.

The emergent phenomenon that physicians opted for the primary action Continued Multitasking 60% of the time versus nurses who opted for the primary action Interruption 62% of the time possibly relates to the type of distractions encountered by each type of clinician or the nature of the CPOE/CIS tasks required by the two types of clinicians. Though data samples were small, physical therapists, occupational therapists, and medical students followed a distraction with an Interruption primary action in all cases. This phenomenon may influence the design of clinician tailored interfaces of CPOE/CIS systems. Further investigation is necessary to determine if the nature of the informatics task, CPOE specific versus CIS specific, influences the secondary action of the clinician and if the type of clinician plays a role in determining the chosen secondary action when encountered with a distracting event.

The taxonomy was able to characterize a consistent rate of Distraction with Multitasking events for physicians using CPOE/CIS during MICU rounds and during non-rounds usage. The nature of the events appeared to differ; yet the frequency of events did not. Analysis of the content of distraction events shows that the rounds events were more structured and clinically focused. The events during non-rounds, in addition to clinically focused communication events, included “casual, polite and social” conversation.

The combined average event rate for the physician observational periods was one event every 3.98 minutes. This event rate shows more frequent occurrence of distraction events than compared to the non-CPOE/CIS specific Coiera et al.’s 2002 study of physicians in an emergency department that detected a distraction event every 11.1 minutes [9].

A similar comparison is shown for the distraction events experienced by nurses in this study to interruptive events experienced by nurses in a level one trauma center in 2005 [19]. Brixey et al.’s definition of an interruption included distracting events and recipient blocked tasks [19] (deferred tasks), allowing for comparison of total events between the studies [12]. In the context of nurse CPOE/CIS use, our study detected an event every 11.75 minutes

compared to the non-CPOE/CIS specific Brixey et al.'s detection of an event every 15 to 24 minutes [19].

The above comparisons to other studies show that general workflow distraction events rates may vary from distraction event rates in the context of CPOE/CIS use. One previous study of distractions in general clinician workflow did detect distraction event rates at 1 every 4 minutes [20], comparable to the CPOE/CIS distraction event rate. However, the internally consistent rates among rounds and non-rounds physician data shows a possible trend of increased distraction event rates in the context of CPOE/CIS use. Our small sample size, yet rich data set, indicates that CPOE/CIS use lends a clinician vulnerable to at least a similar rate and possibly a higher rate of distraction events than are found in general clinician workflow.

Due to the nature of some clinical tasks (i.e. looking up laboratory orders) there exists some difficulty determining the end point of an intended task. Additionally, a secondary action might result in a slip or mistake [16], but we were unable to ascertain if a slip or mistake occurred because we did not have follow-up data [12]. However, the detection of secondary actions, such as Lack of Recall, though a small count, is concerning if the context of CPOE/CIS use does not allow for the recognition and "rescue" of cognitive overload by colleagues at the same level seen in coordinated activity amongst clinicians [4].

Conclusions

Distractions per hour of CPOE/CIS use are as prevalent, and possibly more prevalent, than distractions per hour in general clinician workflow. Evidence of the existence of secondary actions indicates an opportunity for a slip or mistake to occur [17]. The taxonomy is comprehensive enough to capture the distraction events, primary actions and secondary actions that occur in the context of clinicians' use of CPOE/CIS systems.

The interrupt-driven nature of the clinical work environment impacts the cognition of a clinician while using CPOE/CIS. Health care providers rarely work at private workstations in secluded areas; clinical information needs, as well as social engagement, both addressed through clinician communication, contribute to distraction events and possible slips or mistakes [17] in patient care. Given the prevalence of distraction events in the context of CPOE/CIS use, and previous work indicating the relationship between distractions and potential for patient harm, the results of this study indicate an area ripe for further analysis.

Acknowledgments

This research is sponsored by a grant from the National Institute for Nursing Research (T32NR007969) and a grant from the National Library of Medicine (1R01LM07593).

References

[1] Laxmisan A, Hakimzada F, Sayan O, Green R, Zhang J, Patel V. The Multitasking clinician: Decision-making and cognitive demand during and after team handoffs in emergency care. International Journal of Medical Informatics. (in press, 2006.)

[2] Clark H, Brennan S. Grounding in communication. Perspectives on Socially Shared Cognition, Resnick L, Levine JM, Teasley SD, Eds. Washington, DC: Amer. Psychol. Asssoc., 1991, 127-149.

[3] Patterson ES, Cook RI, Woods DD, Render ML. Examining the complexity behind a medication error: generic patterns in communication. IEEE SMC 2004; Part A, 34(6):749 – 756.

[4] Klein G, Woods D, Bradshaw J, Hoffman R, Feltovich P. Ten Challenges for Making Automation a "Team Player" in Joint Human-Agent Activity. IEEE Intelligent Systems. 2004; 19(6):91-95.

[5] Coiera E. Interaction Design Theory. International Journal of Medical Informatics 2003; 69: 205-222.

[6] Van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc. 2006; 13(2):138-47.

[7] Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. JAMIA 2004; 11:104-112.

[8] Alvarez G, Coiera E. Interruptive communication patterns in the intensive care unit ward round. International Journal of Medical Informatics 2005; 74: 791-796.

[9] Coiera E, Jayasuria R, Hardy J, Bannan A, Thorpe M. Communication loads on clinical staff in the emergency department. MJA 2002; 176(9): 415-418.

[10] MEDMARX Data Report 2002.

[11] Bates DW, Teich JM, Lee J, Seger D, Kuperman GJ, Ma'Luf N, Boyle D, Leape L. The impact of computerized physician order entry on medication error prevention. J Am Med Inform Assoc. 1999 Jul-Aug;6(4):313-21.

[12] Collins S, Currie LM, Bakken S, Cimino JJ. Interruptions During the Use of a CPOE System for MICU Rounds. AMIA Symp Proc. 2006.

[13] Flynn EA, Barker KN, Gibson JT, Pearson RE, Berger BA, Smith LA. Impact of interruptions and distractions in dispensing errors in an ambulatory care pharmacy. American Journal of Health-System Pharmacy. 1999; 56(13):1319-25.

[14] Currie LM, Graham M, Allen M, Bakken S, Patel V, Cimino JJ. Clinical information needs in context: an observational study of clinicians while using a clinical information system. AMIA Symp Proc. 2003;190-194.

[15] Chisholm CD, Collison EK, Nelson DR, Cordell WH. Emergency department workplace interruptions: are emergency physicians "interrupt-driven" and "multitasking"? Acad Emerg Med. 2000; 7(11):1239-43.

[16] Zhang J, Patel V, Johnson TR, Shortliffe EH. A cognitive taxonomy of medical errors. Journal of Biomedical Informatics 2004; 37:193–204.

[17] Coiera, E. Communication Systems in Healthcare. Clinical Biochem Review. 2006; 27:89-98.

[18] Brixey J, Robinson D, Johnson C, Johnson T, Turley J, Patel V, Zhang J. Towards a hybrid method to categorize interruptions and activities in healthcare. International Journal of Medical Informatics (in press 2006.)

[19] Brixey J, Robinson D, Tang Z, Johnson T, Zhang J, Turley J. Interruptions in Workflow for RNs in a Level One Trauma Center. AMIA Symp Proc. 2005;86-90.

[20] Spencer R, Logan P. Role-based communication patterns within an emergency department setting. In: Proceedings HIC 2002; 2002; Melbourne; 2002.

Address for correspondence

Sarah A. Collins, RN, BSN
Mail Code 6, 630 West 168th Street, New York, NY 10032
Telephone: 1-781-801-9211, e-mail sac2125@columbia.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Diffusion of Electronic Health Records – Six Years of Empirical Data

Christian Nøhr [a], Stig Kjær Andersen [b], Knut Bernstein [c], Morten Bruun-Rasmussen [c], Søren Vingtoft [c]

[a] *Virtual Centre for Health Informatics, Department of Development and Planning, Aalborg University, Denmark*
[b] *Virtual Centre for Health Informatics, Department of Health Science and Technology, Aalborg University, Denmark*
[c] *MEDIQ, Copenhagen, Denmark*

Abstract

The Danish EHR-Observatory has monitored Danish EHR projects since 2000 with respect to a number of parameters such as diffusion, diffusion rate and the hospital owners expectations. On the basis of the data and a model for technology diffusion three scenarios for future diffusion are built. The results show that the national goal to have EHRs fully implemented in hospitals by 2008 will not be reached. The scenarios built from empirical data provide a useful benchmarking tool for planning and evaluating the EHR implementation programs in hospitals.

Keywords:

Medical Records Systems, Medical Order Entry Systems, state government, diffusion of innovation, technology transfer

Introduction

Several countries follow up on policies to implement electronic health record (EHR) systems for managing and communicating patient information, and a number of them have developed national strategies for the implementation [1-5].

The US does not have a similar national strategy, but on April 27, 2004 President Bush issued an executive order establishing the Office of the National Coordinator for Health Information Technology (ONCHIT) [6].

All the strategies and plans have time estimates for when the EHR systems should be implemented. In the Australian plan they expect that "by 2008, Australia will be well in advance in achieving the goal of electronic connectivity between all major health institutions and health care providers". The Canada Health Infoway goal is "to have an interoperable electronic health record in place across 50per cent of Canada, by population, by the end of 2009". In the US plan the mission is to implement EHR systems nationwide within ten years.

The diffusion of EHR systems among general practitioners in Denmark has been reported earlier [7]. Today practically all GPs (98%) have EHRs in their office, and they communicate 3.5 mill messages every month, where the most common are: 100% of discharge letters, 80% of prescriptions, 98% of lab-reports, and 63% of referrals to hospitals or practicing specialists. The messages comply with the EDI standard and are communicated on the Danish health care network which is developed and administered by MEDCOM. Detailed statistics on current adoption and usage can be found on their homepage [8].

In the Danish revised strategy the goal is to have EHRs fully implemented in hospitals by 2008 [3].

There are however few studies which monitor the implementation status or the actual diffusion of EHR systems. Ford et.al. [9] gathered and synthesized data from six studies on EHR adoption rates. Applying technology diffusion theory they designed a model to project estimated future EHR adoption trends and timelines in three future scenarios, optimistic, best estimate, and conservative. Finally they determined the likelihood of achieving universal EHR adoption. The focus is on physicians in small practices. Their results show that under current conditions, EHR adoption will reach its maximum in 2024, and they conclude that the EHR products now available are unlikely to achieve full diffusion within the 2014 timeframe being targeted by policy makers.

Jha et.al. identified surveys on EHR adoption and addressed their quality [10]. They based their estimates on studies of high or medium quality and found that through 2005, approximately 23.9 percent of physicians used an EHR in the ambulatory settings, while 5 percent of hospitals used computerized order entry systems.

The present study construct a model that project likely EHR adoption patterns based – not on historic estimates as in [9;10] – but on empirical data for the Danish hospitals obtained by the Danish EHR-Observatory through national surveys from 2000 to 2006 [11-16].

The purpose of the EHR-Observatory is to support the realization of the national strategy by monitoring and assess the development, implementation and application patterns of EHR systems in Danish hospitals.

The study makes two new contributions to the EHR discussions and the EHR policy making. First it confirms and substantiates the gap between the optimistic implementation plans seen in several countries and the actual development of useable and accepted systems. Second it provides a benchmarking tool for planning and evaluating EHR implementation programs in hospitals.

Selected for best paper award.

Methods

Identification of what to count

An important limitation in surveying the adoption and diffusion of EHR systems is the definitions of systems handling patient data: Electronic medical record (EMR), electronic patient record (EPR) computerized medical record (CMR), computer-based patient record (CPR), and electronic health record (EHR). There are only minor differences in the meaning depending on the defining country of origin, health sector, professional discipline, and period of time.

There have been many formal definitions of these terms, but they display more similarities than differences with respect to the purpose, functions and goals of electronic records [17].

The International Standard Organization ISO has developed an internationally agreed definition of the EHR: "a repository of information regarding the health of a subject of care, in computer processable form". Here the EHR is defined in terms of its structure, but it is internationally broad and encompasses most of the EHR systems currently used.

These very broad definitions make it difficult to determine the exact object of a diffusion survey. However EHR systems are in general subdivided into components or modules each handling one or more functionalities [18].

The most common components are:

Clinical documentation

Handles progress notes either as free text directly entered into the system or via predefined structured notes. Voice recognition systems are also seen as data entry method.

Physician order entry (POE)

Used for ordering diagnostic test and medication in a standardized and formalized way. Some systems also check for drug interaction and alert for patient allergy.

Booking service

Allow patients and clinicians to book appointments.

Communication/messaging

Facilitates communication between hospitals, General Practitioners, pharmacies, and laboratories.

Results management

Abnormal results warning, trending/graphing.

Charge capture/billing

Coding interventions (this module varies a great deal according to the organization and financing of the health sector).

Decision support

There are a lot of decision support modules in use, but they are rarely used outside the organization where they are developed.

Clinical practice guidelines

Module to manage and maintain clinical guidelines or national reference programs – sometimes categorized as decision support.

Disease management

Management of chronic diseases, like diabetes, etc.

Management of security issues

All EHR systems will have special facilities to manage authentication and authorization of user access according to national legislation

In practice there will be difference between the modules in which a specific functionality is positioned i.e. the above list can neither be regarded as exclusive nor exhaustive.

This is also true for the specific way a work task is executed in practice. E.g. one department can use the booking module to book an x-ray for a patient where another department will use facilities in the CPOE module. This makes it difficult to generate diffusion and adoption rates that can be generalized across different health care systems.

In this study the criteria for having an EHR implemented was a full functioning clinical documentation module, an order entry module able to handle medication functions.

Another limiting factor for the accuracy of surveying diffusion is the unit for measuring adoption. The best choice will depend on the organization and structure of the health sector. In systems where GP's and hospitals work independently, and where the physicians are employed by the hospital with rapidly changing workplaces physicians will not be an adequate unit to count. The number of hospitals will on the other hand be too coarse and the number of specific wards varies so much in size that the count will be inaccurate. In this study the number of beds covered by an EHR system has been chosen.

Questionnaire survey

In the present study the number of beds covered with an EHR system containing functionalities for at least clinical documentation and order entry for medication was surveyed. The ordering of lab tests (clinical chemistry) has already been implemented in all hospitals for a number of years, but is not integrated with medication. The ordering of other tests and X-rays etc. are only partly implemented, but are not included in this survey.

A questionnaire was sent to the county administrations (n=15) every year from 2001 to 2006. The county administrations are the ''owners'' of the public hospitals, and have the political and administrative responsibility for running the hospitals in Denmark. The questionnaire had a kernel of 12 questions that surveyed general aspects of strategy, diffusion, economy, benefits and barriers which remained the same throughout the 6 years and a number of detailed questions to a specific theme of investigation.

The diffusion questions asked the county administrations how many beds were in the hospitals and how many were covered by an EHR system (clinical documentation and medication). Furthermore they were asked how many beds

they were planning to cover the following three years (except in 2001 they were only asked for the current year).

The technology diffusion forecast model

Technology forecasting is a discipline focusing on estimating how technology adoption and diffusion will take place in the future. The goal is not to predict the future, but to create scenarios for planning and policy making. One of the sub-disciplines of technology forecasting is engaged in extrapolating diffusion trends – a quantitative approach to treating adoption data. One of the authors of a technology forecast textbook refers to it as "naive time series analysis" [19]. This implicates that presumptions are met with respect to structure and context of the technology in question. The basic foundation of the essential model however is well proven in countless empirical studies from many areas. The most frequent basic model rests upon the assumption that diffusion of technology follows a sigmoid curve [20]. The sigmoid curve is shown in figure 1 and can be expressed by the equations (1-3):

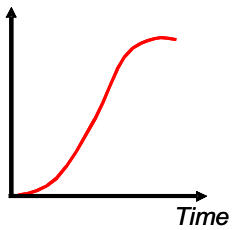


Figure 1 - Sigmoid curve

$$\frac{df}{dt} = b[f/(1-f)] \quad (1)$$

$$f = \frac{100}{1 + c^{*}e^{(-b^{*}t)}} \quad (2)$$

$$f = {}^{t}b[f/(1-f)]dt \quad (3)$$

Where f = the diffusion of the technology in % of the upper limit (here 100), c and b = constants determining the slope and placement of the curve, and t = time. Fisher and Pry have used this expression to study technology forecasting and social change (ref) and is referred to as "Fisher-Pry curve" [21]. If data is available for the first couple of years it is possible to make a best fit curve describing the most likely diffusion course. An optimistic scenario was added by increasing the growth by 20% for the future years, and a conservative scenario was added by reducing the growth in diffusion by 10% for the future years.

Results

The questionnaires were sent to the county administrations electronically each year in April and after one reminder the response was close to optimal as shown in table 1.

The number of beds actually covered and the number of beds the counties expected to cover in the following three years are shown in figure 2. In 2001 we did not ask for the expectations to the following years.

Table 1 - Response rate

Year	Responses	Response rate
2001	12	80,0%
2002	15	100,0%
2003	14	93,3%
2004	15	100,0%
2005	14	93,3%
2006	15	100,0%

The number of beds covered is steadily increasing over the years (the bars boxed with thick lines), but the expectations show a systematically overrating of what turns out to be possible.

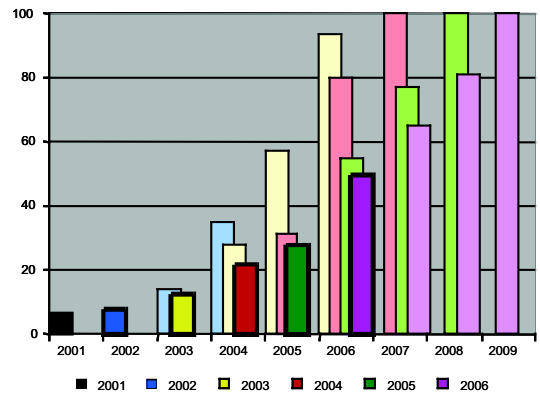


Figure 2 - Actual and expected national EHR diffusion

The data on the actual coverage of beds were fitted into the Fisher-Pry model and an optimistic as well as a conservative scenario was drawn. This is shown in figure 3.

In 2008 the optimistic scenario will have reached 77% adoption, the best estimate will have reached 69% and the conservative scenario 60% adoption.

The optimistic scenario will reach 90% adoption in 2009, the best estimate in 2010 and the conservative scenario in 2011.

Discussion

In figure 2 it is showed how the hospital owners overestimate the capacity to implement the EHR modules. The estimate for the year following the survey is reasonable accurate – only overrated with a few per cent, but the further they predict the implementation the more optimistic they get. This confirm to the optimistic policies on the national level, also seen internationally. In a public financed health care system - to a large extend controlled by politics – there will be a tendency to adhere to the policy goals, even if they seem unrealistic.

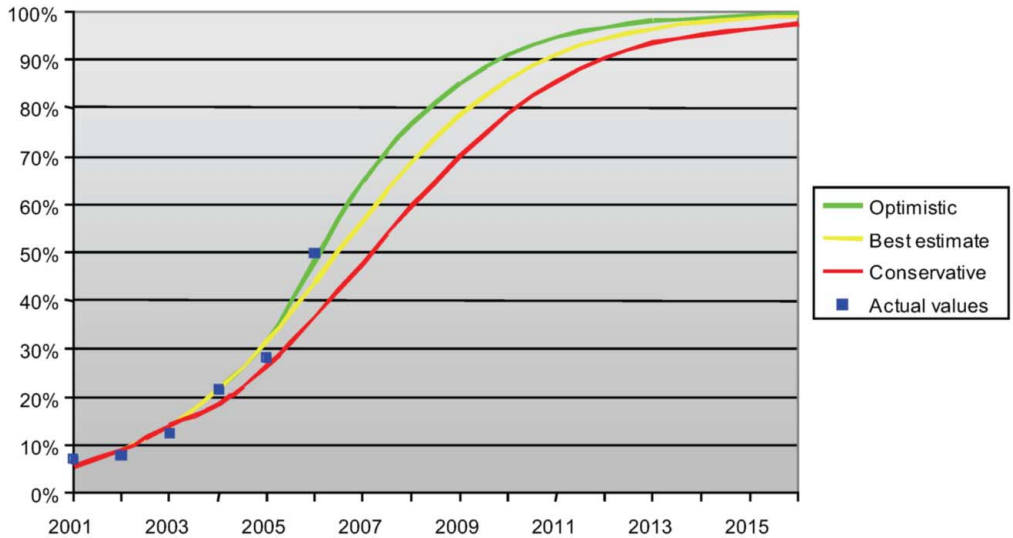


Figure 3 - Three models for diffusion of EHR system based on empirical data from the Danish Hospitals

Another aspect is also that the solutions are not known. There is not a single hospital in the world where a completely flawless system is in place, there are no known places that can be taken as a model. Essential parts have to be developed before usable and acceptable systems that are able to live up to the demands in the national plans exist. The NHS in the UK has a "MythBuster" homepage with the aim to dispel common misconceptions and enable a clearer understanding of the work being done. The answer to the "myth": "*The project is plagued with delays"* is that "*This is because the new software is challenging and there is a debate in the medical professions about what information should go in the records*". [22]

In the conclusion of his book about the history of medical informatics in the United States Morris Collen writes: "Developing a comprehensive medical information system was a more complex task than putting a man on the moon had been," [23]. From a local perspective the development and implementation of EHR systems is simply an unimaginable difficult task.

The scenarios for the diffusion of EHR systems shown in figure 3 are not prognosis. The scenarios are built by a rather static model applying naive time series analysis. Even though there has been numerous empirical studies showing that technology diffusion in general follow the trend in the sigmoid curve, it does not take into account that the investment in EHR systems can be increased by political initiative.

Increase in investment could promote the optimistic scenario, and there are aspects of the diffusion process that can be advanced by investment, such as educating the users to be ready when the systems are introduced, installing all the hardware and networking. For instance to day practically all the Danish hospitals are covered with wireless networks and secure infrastructure – ready for EHR systems, but as in the UK the new software is challenging and the clinical professions, the administrators and the policy makers are still debating the right structure and granularity of the data that must go in to the records.

On the other hand even the rigorous best estimate extrapolation of the diffusion trend might be too optimistic. Jos Aarts has studied implementation of CPOE systems in three Dutch hospitals and three hospitals in the United States. [24] The size of the hospitals he studied was comparable to hospitals in Denmark. The shortest implementation time was at the 315 beds VAPS hospital in Seattle where they replaced a system which had been used for many years. The implementation of the new system took three years. In the other end of the scale is the Radboud hospital in Nijmegen, where they implemented Eclipsys E7000 from 1988 to 2000 – twelve years. It has not been possible to find evidence for any remarkably faster implementations, which indicates that implementation of EHR systems, requires hard work and persistence.

The best estimate extrapolation, and even the optimistic scenario, clearly shows that full implementation by 2008 are unachievable. This initiated a debate about the EHR implementation policy in the beginning of 2006, and resulted in a revision of the policy and the establishment of a central EHR coordination office, which is under manning at the time of writing.

Conclusion

The EHR implementation in Denmark has been monitored by the EHR-Observatory since 2000 obtaining annual data on EHR diffusion. These data has been applied to a technology forecast model, and the results show that the policy for EHR implementation will not reach the goal of full implementation within the appointed time. This usage of technology forecast models can be useful for planning and evaluation of EHR implementation policies in hospitals.

References

[1] Canada Health Infoway. 2003-2004 Annual Report 2004-2005 Corporate Plan Summary. 31-3-2004.

[2] Commonwealth of Australia. HealthConnect Implementation Strategy. 3670 (JN9045). 6-7-2006. Canberra ACT, Commonwealth of Australia.

[3] Danish Ministry of the Interior and Health. National IT Strategy 2003-2007 for the Danish Health Care Service. 2003.

[4] Department of health. Delivering 21st Century IT Support for the NHS - National Strategic Programme. 2002. London, Department of health.

[5] The WAVE Advisory Board to the Director-General of Health. From Strategy to Reality The Working to Add Value Through e-Information (WAVE) Project: Health Information and Management and Technology Plan. 2001.

[6] Bush G W. Executive order: 13335: Incentives for the use of Health Information Technology and Establishing the Position of the National Health Information Technology Coordinator. 2004.

[7] Protti D, Johansen I. Further Lessons from Denmark About Computer Systems in Physician Offices. Electronichealthcare 2003; 2(2):16-21.

[8] MEDCOM. Medcom statistics in English. http://www.medcom.dk/wm109991 . 12-11-2006. 2-12-2006.

[9] Ford EW, Menachemi N, Phillips MT. Predicting the adoption of electronic health records by physicians: when will health care be paperless? J Am Med Inform Assoc 2006; 13(1):106-112.

[10] Jha AK, Ferris TG, Donelan K, DesRoches C, Shields A, Rosenbaum S et al. How common are electronic health records in the United States? A summary of the evidence. Health Aff (Millwood) 2006; 25(6):w496-w507.

[11] Andersen S.K., Nøhr C., Vingtoft S., Bernstein K., Bruun-Rasmussen M. EHR-Observatory Annual report 2002. 2002. Aalborg, EPJ-Observatoriet.

[12] Bruun-Rasmussen M., Bernstein K., Vingtoft S., Andersen S.K., Nøhr C. EHR-Observatory Annual report 2003. 2003. Aalborg, EPJ-Observatoriet.

[13] Nøhr C., Andersen S.K., Vingtoft S., Bruun-Rasmussen M., Bernstein K. EHR-Observatory Annual report 2004. 2004. Aalborg, EPJ-Observatoriet.

[14] Vingtoft S., Bruun-Rasmussen M., Bernstein K., Andersen S.K., Nøhr C. EHR-Observatory Annual report 2005. 2005. Aalborg, EPJ-Observatoriet.

[15] Vingtoft S, Lippert S, Bernstein K, Rasmussen MB, Nøhr C, Andersen SK et al. EHR Observatory Annual Report 2000. 2000. Odense, The County of Funen.

[16] Bernstein K, Rasmussen M.B., Nøhr C., Andersen S.K., Vingtoft S. EHR Observatory. Annual Report 2001. 2001. Odense, The County of Funen.

[17] Schloeffel P. Current EHR Developments: an Australian and International Perspective - Part 1. Health Care and Informatics Review Online™. 2004. 12-8-2005.

[18] Nohr C. Evaluation of electronic health record systems. Methods Inf Med 2006; 45 Suppl 1:107-13.:107-113.

[19] Porter A, Banks J, Mason T W, Rosinni F A, Roper T A. Forecasting and Management of Technology. Wiley-IEEE, 1991.

[20] Rogers E M. Diffusion of Innovations. 5th edition ed. New York: The Free Press, 1995.

[21] Fisher JC, Pry R H. A Simple Substitution Model of Technological Change. Technology Forecasting and Social Change 1971; 3:75-78.

[22] NHS connecting for health. Myth Buster. http://www.connectingforhealth.nhs.uk/faq/mythbusters . 3-12-2006. 3-12-2006.

[23] Collen M F. A History of Medical Informatics in the United States 1950 to 1990. American Medical Informatics Association, 1995.

[24] Aarts J. Understanding Implementation. A sociotechnical appraisal of the introduction of computerized physician order entry systems in Dutch and American Hospitals. Rotterdam: Erasmus University, 2005.

Address for correspondence

Christian Nøhr,
Virtual Centre for health Informatics,
Department of Development and Planning,
Aalborg University,
Fibigerstræde 13, 9220 Aalborg, Denmark
CN@v-chi.dk

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Text Categorization Models for Identifying Unproven Cancer Treatments on the Web

Yin Aphinyanaphongs[a], Constantin Aliferis[abc]

[a]*Department of Biomedical Informatics, Vanderbilt University, Nashville*
[b]*Department of Cancer Biology, Vanderbilt University School of Medicine, Nashville*
[c]*Vanderbilt Ingram Cancer Center, Vanderbilt University, Nashville.*

Abstract

The nature of the internet as a non-peer-reviewed (and largely unregulated) publication medium has allowed wide-spread promotion of inaccurate and unproven medical claims in unprecedented scale. Patients with conditions that are not currently fully treatable are particularly susceptible to unproven and dangerous promises about miracle treatments. In extreme cases, fatal adverse outcomes have been documented. Most commonly, the cost is financial, psychological, and delayed application of imperfect but proven scientific modalities. To help protect patients, who may be desperately ill and thus prone to exploitation, we explored the use of machine learning techniques to identify web pages that make unproven claims. This feasibility study shows that the resulting models can identify web pages that make unproven claims in a fully automatic manner, and substantially better than previous web tools and state-of-the-art search engine technology.

Keywords:

information storage and retrieval, medical informatics, internet, neoplasms, text categorization

Introduction

"The killing of all parasites and their larval stages together with removal of isopropyl alcohol and carcinogens from the patients' lifestyle results in remarkable recovery (from cancer), generally noticeable in less than one week [1]." This is one example of an unproven treatment claim made on the web. These unproven treatments are known as *quackery* with the *quacks* promoting them defined as "untrained people who pretend to be physicians and dispense medical advice and treatment [2]." The internet allows quacks to advocate inaccurate and unproven treatments with documented fatal, adverse outcomes in some situations [3-6].

In regards to cancer patients, Metz et al. reported that 65% of cancer patients searched unproven treatments and 12% purchased unconventional medical therapies online [7]. In another study, Richardson reported that 83% of cancer patients had used at least one unproven treatment [8]. Many patients are ill-equipped to evaluate treatment information [9]. The language and quality of web pages with unproven treatments is also highly variable [10]. The rapid growth of the internet, combined with the ease of publishing unproven claims leads susceptible and often desperately ill patients to further adverse outcomes, patient and family despair, and sunk costs. It is thus an important mandate of the medical profession to protect patients from inaccurate and poor medical information.

So far extensive research has developed several manual methods to combat the propagation of unproven claims on the web. The Health-on-the-Net Foundation advocates self-regulation of health related websites [11]. The foundation applies strict criteria to websites and grants them a seal of approval if they pass. However, most health care consumers ignore the seals [12]. In another approach, experts produced rating tools that consumers are supposed to apply to websites[13, 14]. Another method is manual review of individual websites that are published either in print or electronically.

Each method has limitations. Self-regulation relies on knowledge of the certification and a vigilant public to report failing web sites. Rating tools are dependent on a knowledgeable public to apply, they are difficult to validate, time consuming to produce, and do not always produce consistent ratings [15, 16]. Moreover, the rating tools are not appropriate for use on complementary/ alternative medicine sites [17]. Furthermore, manual review suffers from limits in reviewer time and the selection of web sites to review.

Ideally, we would like a solution that is validated, easy to apply by health care consumers, and works on any webpage. In this paper, we hypothesize that automated approaches to identifying web pages with unproven claims may provide a solution.

Previous work on automatic webpage identification

Previous research focused on automated or semi-automated approaches to identifying *high quality* medical web pages.

Price and Hersh [18] evaluated web page content by combining a score measuring quality proxies for each page. Quality proxies included relevance, credibility, bias, content, currency, and the value of its links. The authors evaluated the algorithm on a small test collection of 48 web pages covering nine medical topics labeled as desirable or undesirable by the investigator. In all cases, the

Selected for best paper award.

score assigned to the desirable pages was higher than the scores assigned to undesirable pages.

Even though the algorithm perfectly discriminated between desirable and undesirable webpages, several limitations exist. The test sample was small and not representative of the scale for a web classification task. The algorithm does not measure content quality directly, but used proxies for quality to compile a score for a web page. The usefulness of some of the explicit criteria may not correlate with content quality [19], and may not be valid or good features to include for scoring.

As a leading search engine, Google has become a de facto standard for identifying and ranking web pages. Pages that rank highly in Google are assumed to be of better quality than those at lower rank. Several researchers have explored this assumption for health pages. Fricke and Fallis [20] evaluated PageRank score as one indicator of quality for 116 web sites about carpal tunnel syndrome. Their results show that PageRank score is not inherently useful for discrimination or helping users to avoid inaccurate or poor information. Of the 70 web sites with high PageRank, 29 of them had inaccurate information.

Griffiths [21] evaluated PageRank scores for depression websites using evidence based quality scores. The authors obtained Google PageRank scores for 24 depression websites from the DMOZ Open Directory Project website. Two health professional raters assigned an evidence based quality score to each site. PageRank scores correlated weakly ($r = 0.61$, $P=0.002$) with the evidence based quality scores.

Tang, Craswell, and Hawking [22] compared Google results with a domain-specific search engine for depression. They found that of a 101 selected queries, Google returned more relevant results, but at the expense of quality. Of the 50 treatment related queries, Google returned 70 pages of which 19 strongly disagreed with the scientific evidence.

Hypothesis

Our fundamental hypothesis for this feasibility study is that we can model expert opinion and build machine learning models that identify web pages that make unproven claims for the treatment of cancer.

To the best of our knowledge, there is no research on validated automated techniques for identifying web pages that make unproven claims. In prior work, we showed that text categorization methods identified high quality content specific articles in internal medicine [23]. Extending this work into the web space, we reverse the hypothesis of the previous studies. Rather than identifying high quality pages, we explore automated identification of low quality pages, specifically pages that make unproven claims for cancer treatment.

Materials and methods

Definitions

Our gold standard relied on selected unproven cancer treatments identified by experts at http://www.quackwatch.org. The website is maintained by a 36 year old nonprofit organization whose mission is to "combat health related frauds, myths, fads, fallacies, and misconduct." The group employs a 152 person scientific and technical advisory board composed of academic and private physicians, dentists, mental health advisors, registered dietitians, podiatrists, veterinarians, and other experts whom review health related claims. By using unproven treatments identified by an oversight organization, we capitalized on an existing high quality review.

Corpus construction

For this feasibility study, we randomly chose 8 unproven treatments from 120 dubious cancer treatments listed by quackwatch.org [24]. The randomly selected treatments were "Cure for all Cancers", "Mistletoe", "Krebiozen", "Metabolic Therapy", "Cellular Health", "ICTH", "Macrobiotic Diet", and "Insulin Potentiation Therapy." We then identified web pages that have these treatments by appending the words "cancer" and "treatment" and querying Google. We retrieved the top 30 results for each unproven treatment. We used a python script to download and store each result as raw html for further labeling.

Corpus labels

We applied a set of criteria for identifying web pages with unproven treatment claims. First, of the initial 240 pages, we excluded not found (404 response code) error pages, no content pages, non-English pages, password-protected pages, pdf pages, redirect pages, and pages where the actual treatment text does not appear in the document[1]. Of the remaining 191 html pages, both authors independently asked the following question of each web page: does the web page make unproven claims about the proposed treatment and its efficacy. We labeled web pages with unproven claims as positive and the others as negative.

Web pages that are purely informational in nature but do not make any unproven claims about the cancer treatment and its efficacy were labeled as negative. Web pages selling a book with user comments that has unproven claims were labeled as positive. Portal pages that do not make any claim were labeled as negative. Web pages that attempted to present an objective viewpoint of the treatment were carefully reviewed for any unproven claims, and, if so, were labeled positive. Additionally web pages that sell unproven treatments but do not make claims were labeled negative.

Both authors applied the criteria independently. We calculated the inter-observer agreement (Cohen's Kappa [25]) at 0.76[2]. Of the 20 sites with discrepant labelings, the reviewers discussed the labels until consensus was reached. The final corpus was composed of 191 web pages with 93 labeled as positive and 98 as negative.

Webpage preparation

For this feasibility study, we chose the simplest web page representation. We converted web pages to a "bag of words"

1 The Google ranking algorithm relies on anchor text to identify web page content. Anchor text may point to a web page that does not use the anchor text in the web page itself.

2 We set a threshold of 0.70 for Cohen's Kappa. If kappa was below 0.70, we would refine the labeling criteria until the threshold was reached.

suitable for the machine learning algorithm[23]. First, for each web page, we removed all content between style and script tags. Second, all tags (including the style and script tags) were removed. Third, we replaced all punctuation with spaces. We split the remaining string on the spaces to obtain individual words. Finally, we stemmed each word [23], applied a stop word list [23], removed any words that appear in less than 3 web pages, and encoded as weighted features using a log frequency with redundancy scheme [23].

Learning model (support vector machines)

We employed Support Vector Machine (SVM) classification algorithms. The SVM's calculate maximal margin hyperplane(s) separating two or more classes of the data. SVMs have had superior text classification performance compared to other methods [23], and this motivated our use of them. We used an SVM classifier implemented in libSVM v2.8 [26] with a polynomial kernel. We optimized the SVM penalty parameter C over the range {0.1, 1, 2, 5, 10} with imbalanced costs applied to each class proportional to the priors in the data [23], and degree d of the polynomial kernel over the range {1, 2, 5}. The ranges of costs and degrees for optimization were chosen based on previous empirical studies [23].

Model selection and performance estimation

We used 10-fold cross-validation that provides unbiased performance estimates of the learning algorithms [23]. This choice for n provided sufficient high-quality positive samples for training in each category and provided sufficient article samples for the classifiers to learn the models. The cross-validation procedure first divided the data randomly into 10 non-overlapping subsets of documents where the proportion of positive and negative documents in the full dataset is preserved for each subset. Next, the following was repeated 10 times: we used one subset of documents for testing (the "original testing set") and the remaining nine subsets for training (the "original training set") of the classifier. The average performance over 10 original testing sets is reported.

To optimize parameters of the SVM algorithms, we used another "nested" loop of cross-validation [23] by splitting each of the 10 original training sets into smaller training sets and validation sets. For each combination of learner parameters, we obtained cross-validation performance and selected the best performing parameters inside this inner loop of cross-validation. We next built a model with the best parameters on the original training set and applied this model to the original testing set. Notice that the final performance estimate obtained by this procedure will be unbiased because each original testing set is used only once to estimate performance of a single model that was built by using training data exclusively.

Quackometer

We compared our algorithm to a heuristic, unvalidated, and unpublished quack detection tool available at http://www.quackometer.net. The exact details of the detection tool are proprietary. In general, the algorithm counts words in web pages that quacks use, and sorts the words into at least 5 dictionaries [27]. It looks for altmed terms such as "homeopathic" and "herbal", pseudoscientific words such as "toxins" and "superfoods", domain specific words such as "energy" and "vibration", skeptical words such as "placebo" and "flawed", and commerce terms such as "products" and "shipping". The algorithm counts the frequency of terms, applies a frequency threshold, and generates a corresponding score from 0 to 10. The tool is available at [28].

We compared our models to the Quackometer by calculating the corresponding area under the curve (AUC) for each 10 fold-split and reporting the mean and standard deviation.

Google PageRank

The PageRank algorithm [29] is used by Google to identify higher quality pages on the web. The basic tenet is that a web page will rank highly if the web page has more and higher quality links pointing to it. For example, if a web page has a link from Yahoo (a highly linked page), it would rank higher than a link from a less linked to web page. In detecting web pages with unproven claims, our assumption is that web pages with poor quality information should get fewer and lower quality links than web pages with better quality.

We used Google as a proxy for PageRank[3]. We made the comparison to our algorithms within each topic rather than within each 10 fold split. We compared within each topic to avoid bias in ranking situations where one topic has uniformly higher Google rank than another topic. We inverted the labels[4] in the 8 randomly selected topics, calculated the AUC, and reported the mean AUC and standard deviation.

Results

Table 1 shows the AUC performance between the machine learning filter models, Quackometer, and Google. The machine learning method identified web pages that make unproven claims with an AUC of 0.93 with a standard deviation of 0.05 across the 10 folds. Quackometer does worse with an AUC of 0.67 and a standard deviation of 0.10 across the same 10 folds. Finally Google performs least effectively in discriminating web pages with an AUC of 0.63 and a standard deviation of 0.17 across the 8 selected topics. Figure 1 shows the corresponding receiver operating curves for each method.

Table 1 – Area Under Curve for Each Discrimination Method

Model	Mean Area Under the Curve
Support Vector Machine	0.93 (std. 0.05)
Quackometer	0.67 (std. 0.10)
Google	0.63 (std. 0.17)*

* The mean and standard deviation are calculated across the 8 topics rather than across the test sets of the 10 folds.

3 Google uses a proprietary version of PageRank for ranking.

4 We test the assumption that PageRank will rank web pages with *proven* claims higher than web pages with *unproven* claims.

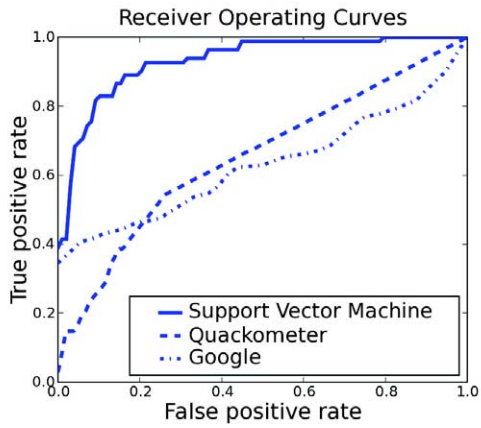


Figure 1: Receiver operating curves for each method.

Discussion

This feasibility study showed that machine learning filter models identify web pages that make unproven claims on a select, focused gold standard. The learning filters have superior performance over the Quackometer [27] and Google. We also note that the loose correlation between Google and high quality sites seems comparable to previous work [20-22].

This method has distinct advantages to rating instruments [13, 14] or manual review. First, there is no need to state explicit rating criteria. The model identified patterns in the data that label a page with unproven claims. Second, compared to the limited focus of manual review on select web pages, these models allow application to any web page.

We also highlight a subtle point in this work. We make a distinction between web pages that make unproven claims and web pages that promote the unproven treatment. Oftentimes, this distinction is blurry. For this work, we only want to identify pages that make unproven claims. Pages that promote a product but do not make unproven claims are not identified. In future studies, we are interested in evaluating models that identify web pages that promote treatments.

In Table 2, we present excerpts from pages where the previous models failed to identify pages with unproven claims. These pages should have been identified by the Quackometer [27] and should not have appeared in the top 30 Google results. Such failure to identify or mark these pages may result in patient's exposure to potentially harmful, unproven treatments.

Table 2 –Web page excerpts where previous tools fail to detect unproven claims. A page that makes unproven claims is identified as such if it has a small support vector machine rank, a large quackometer score, and a large Google rank, respectively. SVM rank is calculated over 10 fold cross validation test set composed of 9 positives and 9 negatives. Google rank is out of the top 30 results returned. Quackometer score provides ranks from 0 to 10. "S" denotes success of the corresponding filter, while "F" failure.

Failure Analysis Excerpts	Support Vector Machine Rank	Quackometer score	Google rank
I am convinced that our mind and emotions are the deciding factor in the cure of cancer.	1 (S)	1 (F)	16 (F)
The hundreds of clinical studies conducted by many competent physicians around the world, including those directed by Dr. Emesto Contreras Rodriguez at the Oasis of Hope Hospital hospital in Mexico, give us complete confidence that there is no danger.	3 (S)	0 (F)	9 (F)
The cure shows results almost immediately and lasts three weeks only. It is cheap and affordable for everybody and proved with 138 case studies.	3 (S)	8 (S)	3 (F)
Many advanced cancer patients are petrified of their tumor. This knee-jerk reaction is caused by orthodox medicine's focus on the highly profitable (and generally worthless) process of shrinking tumors.	1 (S)	1 (F)	18 (F)
IPT (Insulin Potentation Therapy) has an outstanding 135 doctor-year track record (115 years for cancer) over 72 years, and is ready for clinical trials and widespread use.	1 (S)	0 (F)	1 (F)
We are proud of these findings, which confirm that cellular medicine offers solutions for the most critical process in cancer development, the invasion of cancer cells to other organs in the body. Conventional medicine is powerless in this.	2 (S)	1 (F)	8 (F)

In practice, we envision implementing a system that works much like a spam filter works for e-mail. Spam filters identify illegitimate e-mails. In a similar fashion, we envision a system that runs on top of a search engine and flags any web pages that may have unproven health claims.

Limitations

We tested a small sample comprised of 8 unproven treatments in 240 web pages. We will explore how well the models generalize with an independently collected dataset, more unproven treatments, and more labeled web pages. Collecting an independent dataset would allow for validation of the labeling criteria and the model selection procedures.

For this feasibility study, we purposely limited the topic of this study to cancer treatment. In the future, we will build and evaluate other models identifying web pages that make unproven claims for other conditions such as arthritis, autism, and allergies.

Conclusions

We present a first of its kind feasibility study to build machine learning filter models that exhibit high discriminatory performance for identifying web pages with unproven cancer treatment claims. This work paves the way for building broadly applicable models involving more health conditions, more pages with unproven claims, and eventually applied systems to protect patients from quackery.

Acknowledgments

The first author acknowledges support from NLM grant LM007948-02. The second author acknowledges support from grant LM007948-01.

References

[1] Clark H. *The Cure for All Cancers*: New Century Press; 1993.
[2] *American Heritage Dictionary.*
[3] Hainer MI, Tsai N, Komura ST, Chiu CL. Fatal hepatorenal failure associated with hydrazine sulfate. *Ann Intern Med.* 2000 Dec 5;133(11):877-80.
[4] See KA, Lavercombe PS, Dillon J, Ginsberg R. Accidental death from acute selenium poisoning. *Med J Aust*. 2006 Oct 2;185(7):388-9.
[5] Bromley J, Hughes BG, Leong DC, Buckley NA. Life-threatening interaction between complementary medicines. *Ann Pharmacother.* 2005 Sep;39(9):1566-9.
[6] Mularski RA, Grazer RE, Santoni L, Strother JS, Bizovi KE. Treatment advice on the internet leads to a life-threatening adverse reaction: hypotension associated with Niacin overdose. *Clin Toxicol (Phila)*. 2006;44(1):81-4.
[7] Metz JM, Devine P, DeNittis A, Jones H, Hampshire M, Goldwein J, Whittington R. A multi-institutional study of Internet utilization by radiation oncology patients. *Int J Radiat Oncol Biol Phys*. 2003 Jul 15;56(4):1201-5.
[8] Richardson MA, Sanders T, Palmer JL, Greisinger A, Singletary SE. Complementary/alternative medicine use in a comprehensive cancer center and the implications for oncology. *J Clin Oncol*. 2000 Jul;18(13):2505-14.
[9] Sagaram S, Walji M, Bernstam E. Evaluating the prevalence, content and readability of complementary and alternative medicine (CAM) web pages on the internet. *Proc AMIA Symp*. 2002:672-6.
[10] Ernst E, Schmidt K. 'Alternative' cancer cures via the Internet? *Br J Cancer*. 2002 Aug 27;87(5):479-80.
[11] Health on the Net. [accessed 11-27-2006]; http://www.hon.ch/
[12] Eysenbach G, Kohler C. How Do Consumers Search For and Appraise Health Information on the WWW? *BMJ*. 2002 March 9;324.
[13] Bernstam EV, Shelton DM, Walji M, Meric-Bernstam F. Instruments to assess the quality of health information on the World Wide Web. *Int J Med Inform*. 2005 Jan;74(1):13-9.
[14] Kim P, Eng TR, Deering MJ, Maxfield A. Published criteria for evaluating health related web sites: review. *Bmj*. 1999 Mar 6;318(7184):647-9.
[15] Bernstam EV, Sagaram S, Walji M, Johnson CW, Meric-Bernstam F. Usability of quality measures for online health information. *Int J Med Inform*. 2005 Aug;74(7-8):675-83.
[16] Ademiluyi G, Rees CE, Sheard CE. Evaluating the reliability and validity of three tools to assess the quality of health information on the Internet. *Patient Educ Couns*. 2003 Jun;50(2):151-5.
[17] Walji M, Sagaram S, Sagaram D, Meric-Bernstam F, Johnson C, Mirza NQ, Bernstam EV. Efficacy of quality criteria to identify potentially harmful information. *J Med Internet Res*. 2004 Jun 29;6(2):e21.
[18] Price SL, Hersh WR. Filtering Web pages for quality indicators: an empirical approach to finding high quality consumer health information on the World Wide Web. *Proc AMIA Symp*. 1999:911-5.
[19] Fallis D, Fricke M. Indicators of accuracy of consumer health information on the Internet. *J Am Med Inform Assoc*. 2002 Jan-Feb;9(1):73-9.
[20] Fricke M, Fallis D, Jones M, Luszko GM. Consumer health information on the Internet about carpal tunnel syndrome. *Am J Med*. 2005 Feb;118(2):168-74.
[21] Griffiths KM, Tang TT, Hawking D, Christensen H. Automated assessment of the quality of depression websites. *J Med Internet Res*. 2005;7(5):e59.
[22] Tang TT, Craswell N, Hawking D, Griffiths KM, Christensen H. Quality and Relevance of Domain-specific Search: A Case Study in Mental Health. *Info Retr.* 2006;9(2):207-25.
[23] Aphinyanaphongs Y, Tsamardinos I, Statnikov A, Hardin D, Aliferis CF. Text Categorization Models for High Quality Article Retrieval in Internal Medicine. *J Amer Med Inform Assoc*. 2005;12(2):207-16.
[24] Cancer Patients Seeking Alternative Treatments. [accessed 11-26-2006]; http://www.quackwatch.org/00AboutQuackwatch/altseek.html
[25] Cohen J. A coefficient of agreement for nominal scales. *Education and Psych Measurement*. 1960;20(1):37-46.
[26] Chang C, C. L. LIBSVM. 3-13-2006 [accessed; http://www.csie.ntu.edu.tw/~cjlin/libsvm
[27] Science of Quackometrics. [accessed 11-26-2006; http://www.quackometer.net/blog/2006/04/science-of-quackometrics.html
[28] Quackometer. [accessed 11-26-2006]; http://www.quackometer.net/?page=quackometer
[29] Brin S, Page L. The anatomy of a large scale hypertextual Web search engine. *Computer Networks and ISDN Systems*. 1998;30:107-17.

Address for correspondence

Yin Aphinyanaphongs, ping.pong@vanderbilt.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Currency of Online Breast Cancer Information

Funda Meric-Bernstam, M.D.,[1], Muhammad Walji Ph.D..,[2] Smitha Sagaram M.B.B.S., M.S.,[2] Deepak Sagaram M.B.B.S., M.S.,[2] Elmer Bernstam, M.D., M.S.E. [2]

[1]The University of Texas M. D. Anderson Cancer Center and
[2]School of Health Information Sciences, The University of Texas Health Science Center at Houston, Houston, TX

Abstract and Objective

Consumers are increasingly turning to the Web, expecting to find the latest health information. The purpose of this study was to assess the currency of online breast cancer information. We determined whether nine recent advances in breast cancer management were incorporated into 337 unique breast cancer Web pages. Two reviewers independently assessed content; if a Web page covered appropriate advances it was deemed to be "current." Of the 337 Web pages, 89 contained one or more advances. Of the 122 Web pages that had dates of update available, 49% had been updated within 6 months. Only 11%-37% of Web pages covered clinically accepted advances, even among Web pages that were updated after acceptance of the advance into clinical practice. We conclude that online health information is often not sufficiently current. Consumers searching for health information online should always consult an expert clinician before taking action.

Keywords:

breast neoplasms, internet, medical informatics, patient satisfaction

Introduction

The American Cancer Society estimates that over 210,000 women will be diagnosed with breast cancer in 2006. Over 40,000 will die of the disease in the same year (1). Nearly half of women newly diagnosed with breast cancer turn to the Internet for health information (2). The Pew Internet and American Life Project estimates that every day more patients seek health information online than consult physicians (3). The recent US National Cancer Institute-sponsored Health Information National Trends Survey (HINTS) found that health care providers remain a trusted source of health information. However, many respondents went online for health information before going to their physician (4). Consumers are satisfied with their online experience and are making treatment choices based on the information that they encounter (5, 6). Therefore, online information has the potential to significantly impact the health of breast cancer patients.

When breast cancer patients search for health information online, they expect to find the latest advances in medical science. Although there has been appropriate concern regarding the quality of online health information, there are few objective studies of online information currency that, in some cases, is equally important. In this study, we sought to determine whether Web pages that discussed specific breast cancer topics also presented information about important recent advances pertaining to those topics

Methods

Web site selection

We used five popular search engines—Google, Yahoo Directory, AltaVista, Overture, and AllTheWeb—to identify Web sites that consumers are likely to encounter. We used Yahoo Directory because at the time that the study was performed, Google provided the non-directory Yahoo Web search results. For each search engine, we performed 15 searches using the most frequently encountered topics from the first 322 breast cancer-related entries in the NetWellness database of user questions (7) . All searches were performed from June 1 to July 30, 2004. As most Web searchers review only the first page of results (8), we recorded all URLs found on the first search results page, including sponsored (advertisements) and unsponsored results (n = 1585). After we eliminated duplicate URLs, 870 remained.

The URLs were reviewed for relevance (i.e., presence of breast cancer content) by two independent reviewers. If a URL was not relevant, the links on the first page were reviewed and the first relevant link was included in the study. Reviewers were instructed to follow links sequentially, starting from the top of the page, left to right. The reviewers agreed that 344 URLs were relevant. However, each reviewer identified 25 URLs that were unique (e.g., followed different links from the original page). To avoid bias, we included all of these in the analysis, but excluded additional duplicates found at this stage (n=15) (final n = 379).

Web pages were downloaded to a local computer to create a static data set for analysis. Because online content changes continuously, the information stored in search engine databases is usually not an exact copy of the Web in real time. Thus, some sites and pages were unavailable (down) or required redirection and could not be downloaded. Pages were downloaded for two review efforts: the original URL was downloaded for technical quality assess-

ment and the original URL plus pages up to two links away were downloaded for assessment of topics and advances. Downloading the original URL generated four errors, leaving 375 pages. Downloading the page plus pages up to two links away from the original URL generated 30 errors, leaving 349 pages. Six more pages were eliminated at this stage. These included four additional duplicates and two URLs that could not be evaluated for technical quality criteria. One was a survey, not appropriate for analysis and another page was downloaded but without content (empty) and could not be located online. Therefore, a total of 343 pages were evaluated for technical quality criteria as well as breast cancer topics (Table 2, column 2) and advances (Table 2, column 1).

Table 1 - Advances in breast cancer, and number of times they were identified in the survey

Advance	**N**
Sentinel node biopsy*	10
Aromatase inhibitors for adjuvant therapy*	9
Taxanes*	9
Tamoxifen for chemoprevention*	6
Trastuzumab*	5
Minimally invasive breast biopsy*	2
Post-mastectomy radiation*	2
Preoperative chemotherapy for tumor downsizing	2
Partial breast irradiation*	2
Increased public awareness	2
Individualized therapy	2
Estrogen receptor testing for ductal carcinoma in situ	1
Tamoxifen for ductal carcinoma in situ	1
In situ ablation of breast tumors	1
Outpatient treatment	1
Support groups	1
Easier tolerated treatments	1
Demonstration of lack of efficacy of bone marrow transplantation	1
Focused physician training in breast disease	1

* Top nine chosen for further analysis

Identification of advances in breast cancer

A convenience sample of breast oncologists at the University of Texas M. D. Anderson Cancer Center (MDACC) was surveyed to identify advances in breast cancer screening, diagnosis and therapy within the past five years. Free text responses were requested in via email or via personal interviews. Twelve breast oncologists (four surgical oncologists, seven medical oncologists, one radiation oncologist) identified 19 advances (Table 1). Nine specific advances that were identified by more than one oncologist were chosen for further study (Table 2). One advance was related to surgical therapy, three were related to systemic therapy, two related to radiation therapy, one to cancer prevention, and one was related to timing of systemic therapy and surgery.

Results

To determine the currency of online breast cancer information, we identified advances associated with specific breast cancer topics. For example, if a Web page covered breast surgery, it should also discuss sentinel node biopsy, an important advance in surgical therapy. If a Web page covered the appropriate advance it was deemed to be "current." Of the 337 Web pages that covered one or more topics (98% of 343 pages analyzed), 89 (26%) contained one or more corresponding advance. The topics covered and the percentage of Web pages that were current for each advance are shown in Table 2. This was determined as the average of two independent clinically-trained reviewers. Reassuringly, the maximum deviation from the mean for any reviewer score was 2.1%.

Of the 337 Web pages in the study, 124 had listed their date of update. Of these, the date of update on two was not available, due to a technical problem with page fixation. The time between creation of static data set and the last date of update listed for the 122 remaining Web pages are shown in Figure 1. Of the Web pages that listed their date of update, nearly half (49%) had been updated in the past 6 months. The mean time from update was 20 months for commercial (.com) sites, 15 months for organizations (.org/.net) and 14 months for government (.gov) sites. For educational (.edu) sites, the mean time from update was 42 months, but this was based on only two sites.

Once an advance is shown to be effective and introduced as generally accepted or preferred therapy, it should be incorporated into information presented on breast cancer Web pages covering that topic. The date when the advance was considered to be generally accepted was determined as follows. When the advance was a medication, we used dates of approval by the U.S. Food and Drug Administration (FDA) for the corresponding indication. For other interventions, the date of publication of manuscripts describing acceptance of these advances based on consensus conferences or National Comprehensive Cancer Network (NCCN) clinical practice guidelines were used as the date of the advance. Of the advances selected, all but one, partial breast irradiation after breast-conserving therapy, had been determined to be either acceptable practice, or standard of care at the time of the study. Among the 122 Web pages with known dates of updates, we determined the percentage "current" among the Web pages updated

Table 2-Evaluation of content topic and currency

Advance	Topic	Current (%)	Date of Advance	Pages Updated after Date of Advance that are Current (%)
Non-surgical core biopsy	Diagnosis	37%	9/2001 (International Breast Cancer Consensus Conference) (9)	35%
Tamoxifen for chemoprevention	Cancer risk and risk reduction	23%	11/29/1998 (FDA)	27%
Sentinel node biopsy	Surgical therapy	26%	6/30/2002 (Consensus Conference on Sentinel Node Biopsy) (10)	35%
Aromatase inhibitors for adjuvant therapy	Systemic therapy	17%	9/5/2002 (FDA)	29%
Taxanes	Systemic therapy	11%	10/1999 (FDA)	12%
Trastuzumab	Systemic therapy	14%	11/25/1998 (FDA)	16%
Pre-operative chemotherapy for tumor downsizing	Systemic therapy Surgical therapy	13% 16%	11/2000 (NCCN guidelines) (11)	10% 14%
Post-mastectomy radiation	Radiation therapy	37%	9/15/1999 (American College of Therapeutic Radiology and Oncology Consensus statement) (12)	31%
Partial breast irradiation	Radiation therapy	2%	N/A	N/A

after each advance was introduced into standard clinical practice (Table 2).

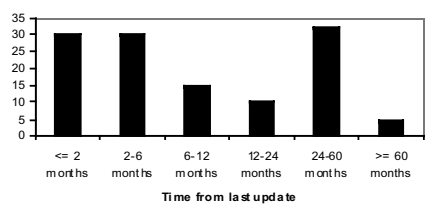


Figure 1 - Time from last date of update

Discussion

Consumers turn to the Web expecting to find state-of-the-art health information. We determined whether nine recent advances in breast cancer management were incorporated into 337 unique Web pages pertaining to those topics. Our data suggest that online breast cancer information is not adequately current, even when pages were recently updated.

Our study is limited by the fact that we focused on one disease state, thus the generalizability of our findings to other topics and search terms, is unknown. However, breast cancer is a common malignancy, and advances in breast cancer management often receive intensive media coverage. Thus, one would expect that online breast cancer information would be especially current.

Only 11-37% of Web pages covered the eight clinical advances that had been determined to be either acceptable practice, or standard of care at the time of the study. All eight advances had been "generally accepted" for at least 22 months prior to our study. It should be noted that the date we selected as "general acceptance" was based on FDA approval or consensus conference, which often lagged by months to years after proof of efficacy by clinical trials. In contrast, oral presentation of a single randomized clinical trial at a national conference in May 1998 was temporally associated with an increase in the use of taxanes for breast cancer in the community; even before study publication in a peer-reviewed journal or FDA

approval in October 1999 (13). This suggests that online information, at least in breast oncology, may lag behind clinical practice.

At the time of the study, the ninth advance, partial breast irradiation, had been shown to be effective in a study conducted at a single institution (14), but was still being tested against the standard of care, whole breast irradiation, in randomized multi-center trials. Only 2% of Web pages that discussed radiation therapy, also discussed partial breast irradiation.

Date of update is a widely recognized Web site quality criterion although it has not yet been shown to correlate with information accuracy (15). In this study, of the 122 Web pages that listed their date of update, nearly half had been updated in the past 6 months. However, the information displayed was "current" in only 10% - 35% of Web pages updated after the medical advances of interest. Thus "currency" based on date of update does not ensure the "currency" of the online information content.

Conclusion

Medicine is evolving rapidly, making it challenging for clinicians themselves to remain up to date. "Current" online health information could provide patients the opportunity to learn about the standard of care and all available options. This would empower the educated consumer to seek out additional options or second opinions in a timely manner. However, our data suggest that online information is not sufficiently current to enable this opportunity.

Acknowledgements

We thank Ms. Marlen Banda, The University of Texas M. D. Anderson Cancer Center, for assistance in data retrieval. We thank the breast oncology group at M. D. Anderson for their clinical input. This study was supported in part by a training fellowship from the Keck Center for Computational and Structural Biology of the Gulf Coast Consortia (NLM Grant No. 5T15LM07093) (M.W.), NLM Grant No. 5K22LM008306 (E.V.B.), a grant from the Robert Wood Johnson Foundation Health-e-Technologies Initiative (E.V.B., F.M-B), and NCRR Grant No. 1UL1RR024148 (E.V.B., F.M-B).

References

[1] Ries LAG HD, Krapcho M, Mariotto A, Miller BA, Feuer EJ, Clegg L, Eisner MP, Horner MJ, Howlader N, Hayat M, Hankey BF, Edwards BK (eds). SEER Cancer Statistics Review, 1975-2003. 2006 [cited 2006 December 2]; Available from: http://seer.cancer.gov/csr/1975_2003/

[2] Satterlund MJ, McCaul KD, Sandgren AK. Information gathering over time by breast cancer patients. J Med Internet Res. 2003 Jul-Sep;5(3):e15.

[3] Fox S, Rainie L. Vital decisions: how internet users decide what information to trust when they or their loved ones are sick. Washington, DC: Pew Internet & American Life Project; 2002 May 22, 2002.

[4] Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med. 2005 Dec 12-26;165(22):2618-24.

[5] Fox S, Rainie L. The online health care revolution: How the Web helps Americans take better care of themselves. Washington DC: Pew Internet and American Life Project: Online; 2000 Novermber 26.

[6] Helft PR, Hlubocky FJ, Gordon EJ, Ratain MJ, Daugherty C. Hope and the media in advanced cancer patients. American Society of Clinical Oncology 36th annual meeting; 2000 2000; New Orleans, LA; 2000.

[7] Ask an Expert. 2003.

[8] Eysenbach G, Kohler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ. 2002 Mar 9;324(7337):573-7.

[9] Image-detected breast cancer: state of the art diagnosis and treatment. International Breast Cancer Consensus Conference. J Am Coll Surg. 2001 Sep;193(3):297-302.

[10] Schwartz GF, Guiliano AE, Veronesi U. Proceeding of the consensus conference of the role of sentinel lymph node biopsy in carcinoma or the breast April 19-22, 2001, Philadelphia, PA, USA. Breast J. 2002 May-Jun;8(3): 124-38.

[11] Carlson RW, Anderson BO, Bensinger W, Cox CE, Davidson NE, Edge SB, et al. NCCN Practice Guidelines for Breast Cancer. Oncology (Williston Park). 2000 Nov;14(11A): 33-49.

[12] Harris JR, Halpin-Murphy P, McNeese M, Mendenhall NP, Morrow M, Robert NJ. Consensus Statement on postmastectomy radiation therapy. Int J Radiat Oncol Biol Phys. 1999 Jul 15;44(5):989-90.

[13] Giordano SH, Duan Z, Kuo YF, Hortobagyi GN, Freeman J, Goodwin JS. Impact of a scientific presentation on community treatment patterns for primary breast cancer. J Natl Cancer Inst. 2006 Mar 15;98(6):382-8.

[14] Vicini FA, Kestin L, Chen P, Benitez P, Goldstein NS, Martinez A. Limited-field radiation therapy in the management of early-stage breast cancer. J Natl Cancer Inst. 2003 Aug 20;95(16):1205-10.

[15] Walji M, Sagaram S, Meric-Bernstam F, Johnson CW, Bernstam EV. Searching for cancer-related information online: unintended retrieval of complementary and alternative medicine information. Int J Med Inform. 2005 Aug;74(7-8):685-93.

Address for correspondence

Elmer V. Bernstam, MD, MSE
Associate Professor
School of Health Information Sciences and Internal Medicine
The University of Texas Health Science Center at Houston
7000 Fannin Street, Suite 600
Houston, TX 77030
Email: Elmer.V.Bernstam@uth.tmc.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Children's Contributions to Designing a Communication Tool for Children with Cancer

Cornelia M. Ruland, RN. Ph.D[ab], Laura Slaughter, Ph.D[a], Justin Starren, Ph.D[c], Torun M. Vatne, MS[a], Elisabeth Y. Moe, Bsc [a]

[a] *Center for Shared Decision Making & Nursing Research, Rikshospitalet Medical Center, Oslo, Norway,*
[b]*Department of Biomedical Informatics, Columbia University, NY*
[c]*Biomedical Informatics Research Center, Marshfield Clinic Research Foundation, Marshfield, WI, USA*

Abstract

In this paper, we describe the roles played as well as contributions made by child participants in the design of an innovative communication tool for children with cancer. SISOM is a handheld, portable computer application with a graphical user interface that is meant to: (1) help children with cancer communicate their symptoms / problems in a child-friendly, age-adjusted manner; and (2) assist clinicians in addressing children's experienced symptoms and problems in patient care. Unlike other applications for children, the purpose of SISOM is not to provide information to ill children but to elicit personal information from them. Thus the application has a unique set of design issues. Healthy and ill children played an important role in different stages in the design process. They made significant contributions to the graphical design of the system's interface; selection of understandable, child-friendly terms used by the system to describe symptoms; iconic and graphical representations; and its usability. We describe the participatory design methods we used that included children and share important insights from this collaborative design process.

Keywords:

participatory design, usability, children, cancer

Introduction

Children diagnosed with cancer are particularly vulnerable. There are many complex physical, functional, psychosocial and behavioral problems associated with having cancer. Children lack life experience and the maturity (both emotional and cognitive) that can equip them to cope with and make sense of their illness. To prevent unnecessary suffering, health care providers need to understand cancer symptoms and problems from a child's perspective. This can be difficult because children vary widely in their illness experiences. Furthermore, children may be prevented from conveying distressing symptoms due to their less developed verbal skills, as well as adults' communication styles and attitude that they should speak on child's behalf. Recent work has shown that clinicians are often unaware of symptoms and problems experienced by pediatric cancer patients [1; 2], which means that many symptoms remain under- diagnosed or untreated.

Children with cancer can benefit from communication support technology that provides them with a "voice"; such a tool can help health care providers understand and address symptoms from the child's perspective. A communication support system for adult cancer patients, called Choice, has been successfully used by patients to report their symptoms, problems and preferences for care. It significantly increased congruence between patient problems and patient care. [3;4] Children with cancer may have even more to gain from similar support, but so far no such systems have been developed for pediatric oncology. Therefore, we developed SISOM, a handheld, portable computer application to: (1) help children with cancer age 7-12 communicate their symptoms/problems in a child-friendly, age-adjusted manner; and (2) assist clinicians in better addressing children's experienced symptoms and problems in patient care.

Developing a tool like SISOM, a communication support system for seriously ill children, is a new area in medical informatics. When seriously ill children are the end users, the challenge is to adapt the application to children's cognitive and emotional developmental stages. SISOM is unique compared to other computer applications developed for children that are mainly computer games or have been developed for informational or educational purposes. These systems primarily deliver information or interactions *to* children. The purpose of most technology designed for ill children so far has been to help them manage and cope with their illness through educational material and play. Systems designed to support patient care have generally focused on obtaining information from parents rather than the sick children. For example, in a recent study Porter et al. [5] built an asthma kiosk within their hospital's emergency department so that parents can provide the critical information and symptoms experienced in order to drive guideline-based care for pediatric asthma. However, no applications have been developed so far that directly elicit and help children communicate their symptom experiences to their care providers as a means to improve patient-centered care.

Selected for best paper award.

Involving children in the design process of an application such as SISOM is critical. Children are experts on being children and can tell us what appeals to them and what they master and understand. When developing SISOM we actively included healthy and ill children in different stages of the design, and children made significant contributions. The design process involved several smaller studies with children as participants. In this paper, we describe the roles children played in the design of SISOM and their contributions to: (1) the graphical user interface (interaction and metaphors); (2) understandable, child-friendly terms that describe symptoms and problems; (3) icons and pictures within the system; and (4) usability. Working with children provided us with important insights that may be helpful for other system developers who wish to design support applications for ill children.

Materials and methods

Children's participation in designing the graphical user interface

Children can play an important role in creating new technologies for other children. Participatory Design (PD) is an essential aspect of good design practice for both adults and children [6]. In our work, we employed PD methods, outlined by Druin [5], that have been successfully applied in the design of story board software, digital libraries for children, computer games as well as educational software [6;7]

PD usually implies that the user participants are as similar as possible as the system's future users. However, when designing an application for seriously ill children, this is only partially possible. PD involves repeated meetings of 1,5-2 hours over an extended period of time [6]. This would pose an impossible burden on children undergoing cancer treatment. When designing SISOM, we therefore, worked together with healthy children, as they still have the aspect of being children in common with our target cancer population. However, ill children participated in other steps of the interface design described later.

Recruitment and sample

For the design sessions and all other methods described in this paper that involved children, IRB approval was obtained. To recruit children to participate in the design of the graphical interface, the Principal of a nearby elementary school in Oslo, Norway, was contacted, asking her permission to send letters to parents of 4th and 6th graders. The letter invited children to participate in the interface design of SISOM, explained the purpose of the system, and outlined the time requirements that PD sessions were 2 hours one afternoon a week, up to 6 weeks in total. Responses were overwhelming. Fifty children (17 boys and 33 girls) replied with interest in participating. As this was far more than needed, we selected final participants based on a phone conversation with their parents, where we screened children based on whether the child: 1) feels comfortable with and is usually active in group activities, 2) uses computers as an educational tool and to play games, 3) is creative, e.g. likes to draw or enjoys building things, and (4) if there was a particular reason why the child wanted to participate. The final group consisted of 12 children who worked in two separate design groups: one group of six 4th graders (9 year olds) and one group of six 6th graders (11 year olds). Other children who had volunteered did participate in other tasks described below.

Design sessions

Design sessions were held at the "Adolescent Club Room" within the pediatric department in Norway's National Hospital, Rikshospitalet, Oslo, Norway, in order to provide a child/adolescent-friendly atmosphere. Both groups attended four sessions after school over a period of two months. Each session lasted two hours. Two of the participatory design sessions were pilot tested with four children of hospital employees (2 boys 2 girls, 9 and 11 years old). Ideas that emerged from the pilot sessions were included in the analysis of design ideas.

Four adults who where trained as "participants" or "observers" were present during each session. Sessions proceeded with the idea that design is an iterative process and a mutual learning experience where both children and adults are experts. The children contributed design ideas, explained which aspects of computer interfaces they found appealing, and gave reasons why some interfaces are better than others. Adults showed the children interface and navigation examples and helped them understand what is technically feasible. Adults also ensured that the system's interface evolved according to a set of system specifications that were defined upfront. Our techniques included role play and scenarios, low-tech prototyping, observation, video-taping and note taking. All groups followed the same procedures.

The first session consisted of three tasks. After introducing each other, and informing the children about the project purpose and plans for the next couple of weeks, children were asked to work in pairs to test four computer programs or games each for about 30 minutes. Each pair had an adult facilitator and an observer who took notes. The purpose was to: (a) learn what interface features and interaction styles children like in computer applications; (b) show children what different features and interactions are possible, in order to stimulate their ideas for SISOM; and (c) as a warm-up. Children were asked to think aloud during this task and comment on what they liked/ disliked about each application, whether it was easy /difficult, engaging/boring and why. They were also asked to describe their favorite computer application and why it was their favorite.

In the next task, children were read a scenario about a child that was sick and had to go to the hospital. Children were asked to tell us and role play what these symptoms felt like. We choose a stomach flu scenario rather than a cancer story because we anticipated that children were more likely to have had a real life experience with flu rather than with cancer. We also felt that having a detailed scenario dealing with cancer might make the children worry or frighten them. However, children were asked to role play symptoms that are contained in SISOM, such as nausea, fatigue, and feeling sad. The story ended with a nurse com-

ing to the hospitalized child, handing her a tablet computer that the child could use to report its symptoms, problems and worries, so the doctor and nurse could help.

The children were then asked to start designing a system based on the scenario they had just heard about, and draw or build what it could look like and how it would behave. Children were provided with a large table equipped with low-tech prototype material, including paper in different colors, crayons, pencils, cardboard, cartoons, lego, glue etc. and asked to work on the task in groups of 2-3. Towards the end of each session, each sub-group was asked to summarize their ideas so that the rest of the group could provide feedback.

In sessions two-four, the team continued to draw and discuss different aspects of the interface, either in pairs or as a whole group. All sessions were videotaped, and notes were taken by adult observers.

Figure 1 - Drawings made during PD sessions

We made particular efforts to help our "healthy participants" understand the purpose and context of SISOM since they were not our real end-users. Therefore, all sessions started with a story of a sick child that the children could recognize or easily imagine. Additional symptoms and problems were introduced, again from the SISOM symptom list. In the third session, the children role-played an actual hospital scene where they took turns being dressed in a hospital gown, and put into bed with a pretend I.V. line. The other children played the role of doctor, nurse or worried parent. The "sick" and bandaged child in bed tested different computer types and input devices and was asked for feedback. They also tested the adult application for cancer symptom assessment that has been successfully used by cancer patients over the last years. Children found the adult version relatively easy-to-use, but naturally found its interface boring because symptoms are only presented as text in tables.

Between sessions, the adult team discussed ideas and observations from the previous session, monitored progress, summarized what was learned, and adjusted, when necessary, the agenda for the following session. The graphical designer on the team drew out children's ideas in Adobe Illustrator that were given back to the children for further evaluation and elaboration in the next session. Figure 1 shows an example drawn by two 11 year old girls that depicts their idea of how SISOM could help children report where it hurts. The left side displays the children's drawing; the right side how this idea was refined. In the final version the figure can be displayed as a boy or a girl, dependent on the gender of the child user, and it can also be turned to show the posterior side.

Table 1 – Children's Design Contributions

Idea Category	**Count**
Age /gender specific user choices	4
Animations, colors, graphics, media	19
Background scene	11
Interactions	30
Sound, text, voice	27
Main characters, avatar	16
Navigation	27
Help functions	11
Input	10
Likes, dislikes, experience	6
All	161

The final data from the design sessions consisted of the low tech prototypes that were developed, observation notes, and 22 hours of videotape. Videos and observation notes were annotated and organized according to topic, such as ideas for background, help, navigation, animation, graphics, age and gender specific content, as well as additional material. The children came up with a total of 161 design ideas altogether, but some of the ideas were the same or very similar; we counted a total of 109 (68% of 161) unique ideas. Children's idea types are summarized in table 1.

Children had many excellent and creative suggestions that the design team would not have thought of. For example, one design idea that emerged from the sessions and was endorsed by both boys and girls in both age groups was a "sailing from island-to-island" navigation theme. The children suggested that symptoms be placed on islands and that the sick children can sail from place to place in order to select symptoms. Also, one idea was that each child could create a main character and select its appearance (hair, cloth etc). However, not all ideas were workable. For example, when designing an interface for younger children who cannot read, one group suggested that a voice could be activated by clicking a button where: "Click here if you cannot read" is written on it. One group of boys suggested

shooting at the body picture with a gun to mark where it hurts.

Children's contributions to child-friendly terms

The participatory design sessions were preceded by the development of the list of symptoms to be contained in SISOM. This list was based on a critical review of 98 articles from the scientific literature that addressed symptoms and problems encountered by children with cancer along physical, functional, psychosocial and behavioral dimensions. In addition, we conducted focus groups with clinical specialists (physicians, nurses, psychologists, social workers) and parents who critically reviewed the symptoms abstracted from the literature and supplemented them with expert opinion. They were also asked what terms and expressions children used when communicating about symptoms with them. This resulted in a preliminary list of 78 child-friendly symptoms and problem expressions.

To ensure that children could understand the symptom and problem terms contained in SISOM, we interviewed 14 children (7 healthy volunteers, 5 with cancer, age 8-12). We presented them with one symptom at a time in random order and asked them about: (a) their understanding of the term: "what does it mean to be… (e.g. nauseated?)"; (b) alternative expressions: "how would *you* say it to let others know that you were…?"

Interviews were audio-taped, transcribed and analyzed. Symptom/problem terms were coded into: (a) clearly understandable terms; (b) ambiguous terms; (c) does not know what the term means; (d) familiar with the term but assigns a different meaning to it. Terms that were not clearly understandable were revised. Children provided several excellent suggestions for meaningful child-friendly terms that were used in these revisions.

Evaluations of graphical representations

PD sessions provided us with design ideas that were implemented in a prototype that subsequently had to be validated. Graphical elements function best when they are metaphorically meaningful and correctly represent the concept in question [8]. In SISOM each symptom is represented with a picture. Therefore, we conducted evaluation sessions with five children from our volunteer group age 9-11, to learn if they could correctly recognize the symptoms based on the graphical representations. Children were presented with one symptom picture at a time without providing them with any labels, or category headings under which they belonged. This approach evaluated the pictures only; in SISOM, symptoms are displayed in a graphical background that provides the context and text labels that can be heard by clicking on an icon if the child cannot read. Figure 2 displays a picture of the symptom "difficulty sleeping" that was immediately recognized by all children.

Figure 2 – Difficulty sleeping

A facilitator conducted the interviews and an observer took notes. Sessions lasted on average 45 minutes. For each picture, children were asked what symptom they thought it depicted and why. Their explanations helped us identify cues in the picture that made the symptom recognizable. If children did not recognize the symptom, we provided them with the category name under which it belonged. Children were provided with coloured pencils to revise the pictures or draw new ones.

Pictures were grouped into four categories: A: the symptom was correctly recognized immediately; B: was recognized after knowing the category name under which it belongs; C: knowing the symptom, the picture is a good representation; D: the picture does not represent the symptom well. 61% of the symptoms were immediately recognized by all children. 12% of the pictures were considered not a good symptom representation and were revised. Children provided a number of excellent ideas for revisions that we used.

Children as usability testers

We performed several usability tests with healthy and sick children at different stages of the development process. Usability testing encompasses a range of methods for identifying how users actually interact with a prototype or a complete system. It is an iterative process that involves testing the system and then using the test results to change it to better meet users' needs. The best process is to try out a prototype with a few users, fix it, and test it again. [9]

At the current stage of SISOM's development, four healthy school kids and two ill children at the hospital have participated in usability testing. They were asked to use the system and select a set of symptoms. They were prompted to think aloud during the task which was videotaped. Morae™ software was used for automatic recording and analysis of all events on the screen, such as when the user clicked on an object, opened a dialog box, or viewed /listened to specific text.

Usability testing showed that children in our target age group are very computer savvy, and have no fear of clicking on and finding things intuitively. Few initial instructions sufficed. All children liked the ability to select the main figure, and the metaphor of sailing through the island world provided them with a sense of discovery.

However, we made the interesting observation that a system can evoke very different associations in children with the experience of a life threatening illness than in healthy children. For example, in an early prototype, when the user was done with a symptom and zoomed out of the picture a spark appeared, the main character disappeared, and did not re-appear before the child had clicked on another picture. One of the sick children thought that the main figure exploded (= died). This made him very uncomfortable. Thus transitions between pictures were subsequently changed. The healthy kids had no such associations and thought that the spark feature was cool. It became clear from this experience that sick children who are the end-users are important usability testers, and healthy children cannot serve as their proxies in this task.

Discussion

We can summarize several important observations from working with children in the design of SISOM. In our project, children were involved in a number of tasks and made significant contributions to the system's graphical interface, child-friendly terms, iconic and graphical representations, and usability. Based on the tasks we gave them, children were able to contribute very useful ideas that the adult designers would not have thought of and considerably improved the software. It was also crucial that they participated in evaluations of the system.

However, children have also limitations. They are not professional designers and do not always have a good grasp of logical design. Good design for ill children requires insights and knowledge that children do not have (e.g. expert knowledge of children's motor and cognitive abilities, pedagogy, and child psychology). We had to make sure to meet the goals of SISOM and a set of pre-defined criteria. Specially, we had to ensure that the software allows health care providers to obtain valid and reliable data from children about their symptoms in a clinical context, without being too time-consuming and challenging. Our child participants often focused on designing fun and often time-consuming aspects, such as funny noises along with vivid animations for a symptom such as throwing up. In spite of reminders that we were not designing a game, the children had the tendency to slip back into a "game mode". Also, children could spend considerable time on fine details and lose sight of the overall purpose. For example, they could spend a whole session on drawing flowers in a background landscape or on ways to choose eye and hair color for a figure created to navigate the system. Thus the knowledge of professional designers is crucial when it comes to final decisions about which ideas to implement.

Comparing the contributions of 9-year olds to 11 year olds, we did not see any significant differences. Both did well in performing the tasks and suggested creative design ideas that could be used with children in their own age group. Working with two age groups was very valuable. Older children in our group were very sensitive to not creating a system that appeared childish. Smaller children had difficulty coming up with ideas suited for an older age group. Boys suggested a lot of action, guns, rockets, a race course, shooting etc. While not completely free of action, girls' ideas contained more flowers, animals, and "softer" effects. This observation suggests that it may be worthwhile to not only offer age-adjusted interface choices, but also gender-adjusted options.

A challenge in the design of SISOM was to balance participation of healthy and ill children. The burden and time required for some of the tasks prevented us from asking sick children to participate, e.g. in design sessions focusing on the graphical user interface. This however, raises the question whether healthy children can conceptualize what it is like to be suffering from a serious illness, and thus the degree to which they can serve as proxies in participatory design and evaluations. Our experience is that this is only partially possible. Role play and scenarios certainly helped to increase healthy children's understanding of what SISOM was for. Still, it was difficult for them to grasp the context in full. We made however an interesting observation. One of the 9-year olds had a cousin with lymphoma and thus had been exposed to cancer in her family. While we have no other empirical data to support this, this child appeared to have a more mature grasp of the purpose of the system. Her knowledge of some aspects of cancer, through her life experiences with her cousin, caused her to have more focused ideas based on a more sophisticated knowledge. Thus it seems that personal experience is an important factor for valuable design contributions. The limitation of healthy children may be partially compensated through extensive usability testing where participation of ill children as end-users is imperative. This will allow us to discover weaknesses in the design where sick children can contribute ideas for improvement.

Acknowledgements

This study was supported by grants from the Norwegian Research Council, The Norwegian Nurses' Association, Rikshospitalet's research fund, and Medinnova. For their valuable contributions we thank: Roar Andersen, Yizhak Itzchaki, Denis Pototylyuk, Birgitte Kolbeinsen, Simon Hjelle, and Glenys Hamilton.

References

[1] Patenaude AF, Last B: Cancer and children: where are we coming from? Where are we going? Psycho-Oncology 2001;*10:*281-283.

[2] Camp, LD: A comparison of nurses' recorded assessments of pain with perceptions of pain as described by cancer patients. Cancer Nursing 1988;*11:*237-243.

[3] Ruland CM. Handheld Technology to Improve Patient Care: valuating a Support System for Preference-based care Planning at the Bedside. *.Am.Med.Inform.Assoc.* 2002;9(2):192-201.

[4] Ruland CM, White T, Stevens M, Fanciullo G, Khilani SM. Effects of computerized system to support shared decision

making in symptom management of cancer patients: preliminary results. JAMIA 2003; 10(6):573-579.
[5] Druin A: The Design of Children's Technology. San Francisco CA, Morgan Kaufmann Publishers, 1999.
[6] Scaife M, Rogers Y, Aldrich F, Davies M: Designing for or designing with? Informant design for interactive learning environments; Proceedings of CHI '97, ACM Press, 1997, pp 343-350.
[7] Marcus A. Icons, Symbols and Signs: Visible languages to facilitate communication, 2003.
[8] National Cancer Institute, United States National Institutes of Health, Bethesda, Md Usability Website. http://www.usability.gov/methods/usabilitytesting.html
[9] Porter SC, Kohane IS, Goldmann DA: Parents as Partners in Obtaining the Medication History. JAMIA 2005;*12:* 299-305.

Address for correspondence

Cornelia Ruland, RN, Ph.D
Center for Shared Decision Making and Nursing Research
Rikshospitalet-Radiumhospitalet HF Forskningsvn. 2b
0027 Oslo, Norway.
E-mail: cornelia.ruland@rikshospitalet.no

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

"It's Your Game": An Innovative Multimedia Virtual World to Prevent HIV/STI and Pregnancy in Middle School Youth

Ross Shegog, Christine Markham, Melissa Peskin, Monica Dancel, Charlie Coton, Susan Tortolero

Center for Health promotion and Prevention Research, University of Texas School of Public Health, Houston, TX, USA

Abstract

Early sexual initiation is associated with increased risk of unintended pregnancy and sexually transmitted infection (STI). Effective HIV/STI/pregnancy prevention interventions for middle school youth are urgently needed. "It's Your Game, Keep It Real" (IYG) is a curriculum delivered in 7th and 8th grade that combines classroom activities with individualized, tailored computer-based activities embedded in a 'virtual world' environment. Interactive multimedia can offer a confidential, tailored, and motivational educational experience. Virtual world game interfaces offer further potential to immerse the learner. The purpose of this study was to evaluate the multimedia education program component of IYG on student attitudes of importance of the curriculum content, self-efficacy regarding refusal skills, and usability parameters of ease of use, credibility, understandability, acceptability, and motivation to determine that a broader efficacy field test would be indicated. Results of the study indicated acceptable usability criteria and impact on short-term psychosocial outcomes. IYG is currently being evaluated in a randomized controlled trial in ten Texas middle schools.

Keywords:

HIV/pregnancy prevention, multimedia, youth

Introduction

Although adolescent pregnancy and birth rates have declined steadily in the United States over the past decade, adolescent pregnancy remains a serious public health issue.[1] In Texas, 56,086 births reported in 2000 were to mothers aged 10-19; 8,465 of these were reported in Harris County.[2] In 2001, Texas also ranked joint first among states for rate of repeat births to teen mothers: 25% of births to mothers aged 15-19 were repeat births.[3] Teen pregnancy is particularly prevalent among minority populations, especially African American and Hispanics. [2,4] For the teen mother, consequences of pregnancy include decreased likelihood of completing high school and increased likelihood of relying on welfare. Infants born to teen mothers have lower birth weights, are more likely to perform poorly in school, and are at greater risk of abuse and neglect.[5] Overall, the U.S. government spends over $25 billion a year for social, health, and welfare services to families begun by teen mothers.[5]

Adolescent sexually transmitted infections (STI) including HIV also represent serious public health problems. In 2000, youth between the ages of 15 and 24 accounted for 9.1 million (48%) of all new STI cases.[6] The estimated medical cost of these cases was $6.5 billion.[7] At least half of all new HIV infections are estimated to be among those under the age of 25 and most young people are infected through sex.[8] Among youth, teen girls and minorities have been particularly affected.[9] Texas currently ranks fourth among states for the estimated number of persons living with HIV/AIDS.[10,11] In Houston, 319 (75%) of HIV cases reported between 1999-2003 among 13-19 year olds were among African American youth and 59 (13.9%) were among Hispanic youth.[11]

According to data from the 2003 Middle School Youth Risk Behavior Surveillance Survey (MSYRBS), 16% of 7th graders and 19% of 8th graders have engaged in sexual intercourse (implicitly, vaginal intercourse).[12] This is disturbing because early initiation of sexual intercourse has been associated with an increased risk of STIs and pregnancy.[13-16] Among older adolescents, condom rates for oral and anal intercourse are typically lower than for vaginal intercourse,[17-19] and a history of anal sex is predictive of non-use of condoms during vaginal sex.[20] This evidence points to the urgent need for effective HIV, STI and pregnancy prevention interventions at the middle school level to help delay or mitigate the consequences of early sexual activity.

A 2004 review of HIV, STI and pregnancy prevention programs for middle school populations identified 12 programs that have been rigorously evaluated.[21] Of these programs, seven showed positive behavior change. Two recently published studies also reported behavioral change for interventions at short-term (5 month)[22] and twelve-month follow-up.[23] These programs all incorporated small group activities to address peer pressure and decision-making, role-playing to practice refusal and communication skills, and other interactive, experiential learning techniques. Although these programs all demonstrated some positive behavior change, most showed

Selected for best paper award.

differential impact by gender or by sexual experience.[24-26]

The "It's Your Game" (IYG) curriculum offers an innovative application of computer-based gaming technology, interactive computer-based activities, and small group classroom interaction that does not currently exist in middle school HIV/STD prevention curricula. The purpose of this study was to evaluate a component of the curriculum, the IYG virtual world multimedia education program, to determine its impact on short-term outcomes and usability parameters to ensure feasibility of the program for field testing in a 10-site randomized controlled trial.

Methods

Study Design: The design was a single group pre-test post-test usability study conducted in facilities at the University of Texas, Houston Health Science Center.

Subjects: A convenience sample (n=14) of Houston middle school students was recruited who were representative of the Houston Independent School District demographic: primarily minority (50% African American) and of approximately equal gender (57% female). This sample tested the 7th grade lessons. Nine students from this sample returned 7 months later to test the 8th grade lessons. This 8th grade sample was 12-14 years of age, 55% African American, and 55% female. A smaller sample size, used in this study, was consistent with recommendations for usability testing given that statistical significance is not required to determine major usability problems and that the best cost-benefit ratio is achieved with 3-5 users in each representative group.[27] Participation was voluntary and written parental consent and child assent was obtained.

Intervention: The conceptual framework for the IYG curriculum is based on Social Cognitive Theory (SCT), social influence models and the Theory of Triadic Influence (TTI).[28-30] The curriculum consists of 12 lessons delivered in 7th grade and 12 lessons delivered in 8th grade. In each grade, the curriculum integrates group-based classroom activities (e.g., role plays, group discussion, and small group activities) with personalized journaling, and the IYG virtual world multimedia intervention delivered on laptop computers. The virtual world intervention comprises a total of 8 lessons spread through the 7th and 8th grade curriculum. These lessons comprise 3 functional elements: (1) A 3D virtual world interface featuring an entertainment complex motif (Figure 1); (2) tailored educational activities including interactive 2D exercises, quizzes, animations, peer video, and fact sheets that target determinants of sexual risk-taking (Figure 2); and (3) "real world"-style teen serials with on-line student feedback which allow for real time group discussion in the classroom. In addition, selected computer activities are tailored by gender or by sexual experience and intent so that students receive information and skills-training that is tailored to their needs.

Figures 1 & 2 - Screen captures of an aspect of the virtual world and an interactive FLASH activity

A life skills decision-making paradigm (*Select, Detect, Protect*) underlies the activities, teaching students to *select* personal limits or rules regarding risk behaviors, to *detect* signs or situations that might challenge these limits, and to use refusal skills and other tactics to *protect* these limits. Specific topics covered in the 7th grade include characteristics of healthy friendships, setting personal limits and practicing refusal skills in a general context (e.g., regarding alcohol and drug use, skipping school, cheating), information about puberty, reproduction and STIs, and setting personal limits and practicing refusal skills related to sexual behavior. The 8th grade curriculum reviews these topics and presents additional activities regarding the characteristics of healthy dating relationships including age compatibility, the importance of HIV, STI, and pregnancy testing if a person is sexually active, and skills training regarding condom and contraceptive use.

Study Protocol: The Virtual world intervention was tested separately from other classroom curriculum elements. The students accessed the program in a simulated class setting. Each student was provided with a laptop computer with headphones and asked to complete each of the four 35 minute 7th grade computer lessons individually. At the end of each lesson each student completed feedback questionnaires. At the end of the 4th (final) 7th grade lesson each student also completed overall assessment of the program. Sessions were observed by study personnel who logged problems (technical or content related) and provided assistance as required. This protocol was repeated 7 months later for the 8th grade lessons by the students, who were then in 8th grade.

Data Collection: Questionnaire, computer-based, and observational data was collected. Demographic data included age, race, gender, computer experience, type and frequency of computer use. Attitudes to the use of computer-assisted instruction were collected using a validated questionnaire.[31] Pre and post lesson ratings on the importance of content domains and self-efficacy (confidence) in performing skills was assessed using semantic differential scales developed for the intervention and embedded in the virtual world interface. Usability parameters including ease of use, credibility, understandability, acceptability, and motivation were assessed using Likert scale ratings adapted from usability assessment instruments reported by Sapperstein et al, 2004.[32] Open-ended

responses on recommendations for improvement of the program were collected via computer and paper and pencil questionnaire.

Data Analysis: Data was analyzed using descriptive and inferential statistics (paired t-tests) with SPSS analytic software.

Results

Demographics: The students were experienced computer users. Most accessed computers at their home (89%), friends (55%), and school (50%). Thirty three percent used computers for more than 1 hour on week days and 77% for more than one hour on weekends. Most frequent uses of computers were school work, visiting websites, and e-mail (all > 50%). Approximately 44% of the sample reported achieving A or B grades at school.

Short-term outcomes: Student attitudes toward the use of computers in education was enhanced in the 7th grade sample, and significantly enhanced in the 8th grade sample ($p<0.5$). Ratings of the importance of the program content in each lesson increased significantly for all content domains including the importance of keeping 'good' friendships, understanding how reproduction works and the possible consequence of sex (HIV/STDs/pregnancy), and the importance of enacting behaviors to limit sexual experience, described as selecting, detecting, and protecting personal rules about choosing not to have sex (all $p<0.05$). Ratings of self-efficacy for enacting behaviors in these domains also significantly improved (all $p<0.5$).

Usability: Usability parameters were highly rated across 7th and 8th grade lessons. *Ease of use*: Virtual world and educational activities were rated as easy to use by 78-100% of students. *Credibility:* A minimum of 92.9% of students perceived the content as correct and trustworthy across all lessons. *Understandability:* There was 100% agreement that most words in the program were understandable. *Acceptability:* Most students (71-100%) rated interface strategies and specific program activities within each lesson as fun and 92.9-100% rated it as helping them make healthy choices regarding sexuality. There was 64.3 – 92.9% agreement that each of the program lessons were as much or more fun than other lessons or favorite video games and 85.7-100% of students rated the lessons as "just right". Open ended responses suggested satisfaction with the lessons with just two participants suggesting a desire for more media elements in the form of characters and movies.

Discussion

The "It's Your Game" virtual world multimedia HIV/STD and pregnancy prevention lessons were found to significantly enhance attitudes toward the importance of sexual risk behavior and self-efficacy toward initiating such behavior. The lessons were reported to be easy to use, credible, understandable, acceptable, and of sufficient motivational appeal to students (12-14 years of age) to elicit confidence that the program was feasible for field testing in middle schools as a component of the IYG curriculum. An unexpected 'by-product' of this study was the immediate positive effects on attitudes toward the use of technology in education. Together, these results tender support for the emerging interest in gaming interfaces through burgeoning Serious Games and Games for Health research initiatives.[33,34] The capability of computer-based applications to provide interactive and individual tailored experiences that are also confidential is particularly salient when considering the sensitive nature of HIV/STD, pregnancy, and sexual health program content. Further, the technology can provide views and experiences beyond the limited classroom forum, allowing students to gain a realistic appraisal of what constitutes normative sexual behavior. A strength of the program is its ability to provide anonymous input that can be subsequently used within the small group setting of the classroom.

While the small sample size and limited scope is appropriate for the objectives of a usability study of this type [27], some results needs to be viewed with caution. A single group pre- post-test study design was the basis for the reported change in importance, self-efficacy, and attitude to computer-assisted instruction. While these results are encouraging, this design is open to internal threats to validity and does not imply causality.[35] Also, these effects were measured immediately prior to and following the intervention with no long term follow-up.

The combined IYG curriculum is currently being evaluated in a randomized controlled trial in 10 Texas middle schools. Baseline data have been collected from a largely minority sample of 1,321 7th graders comprising 57.1% female, 43.5% black, and 41.9% Hispanic, with a mean age 12.5 (SD=0.69) years. Fourteen percent had engaged in any type of intercourse (12% vaginal, 7.9% oral, 6.5% anal). Impact and outcome data are being collected 5, 14 and 24 months post-baseline. Measures include sexual behavior and intentions (lifetime/current vaginal/oral/anal sex, condom use), beliefs, perceived norms, knowledge, self-efficacy, reasons to/not to have sex, exposure to risky situations. This study will provide behavioral outcome data for the curriculum. Initial 5 month post-test results indicate students receiving the IYG curriculum report program lower prevalence of any type of intercourse during the past 3 months, and positive change in abstinence beliefs, perceived friends' beliefs about sex, exposure to risky situations, and reasons not to have sex (all $p<0.05$).

If the IYG curriculum proves effective, dissemination strategies will include a web-based application, a Spanish language version, and "institutionalization of the curriculum within the Houston Independent School District. Future development plans for the IYG virtual world multimedia program are to evaluate its impact when used as a stand alone intervention, separated from the IYG curriculum. The potential for the program to provide multi-user education for naturally occurring peer groups is also being investigated.

Conclusion

These results indicate that the “It’s Your Game” virtual world computer-based program is a feasible modality for HIV/STD, and pregnancy prevention for middle school students. The program meets usability criteria and shows potential in impacting attitudes and self-efficacy regarding refusal skills. The program also shows promise in translating positive attitudes to the use of technology in education. This usability study indicated that field testing on behavioral effects was indicated. The IYG curriculum, containing the IYG virtual world multimedia program, is currently being evaluated in an NIMH NIH funded randomized controlled trial in 10 Texas middle schools to determine its impact on sexual behaviors and their antecedents.

Acknowledgement

This study was funded by the national Institutes of Mental Health, NIH. Award # NIMH R01 MH66640.

References

[1] Klein JD, Barratt MS, Blythe MJ, Diaz A, Rosen DS, Wibbelsman CJ, Hertweck SP, Kaufman M, Shain B. Adolescent pregnancy: Current trends and issues. Pediatrics. 2005;116(1):281-286.

[2] Texas Department of Health Bureau of Vital Statistics. 2000 Birth Report: Births by Mother's Age and Ethnicity. Data geocoded by the Center for Health Policy Studies, UT-Houston, School of Public Health, 2000.

[3] Texas Department of Health Bureau of Vital Statistics. 2000 Birth Report: Births by Mother's Age and Ethnicity. Data geocoded by the Center for Health Policy Studies, UT-Houston, School of Public Health, 2000.

[4] Alan Guttmacher Institute. U.S. Teenage Pregnancy Statistics Overall Trends, Trends by Race and Ethnicity and State-by-State Information. Alan Guttmacher Institute. Feb 19, 2004. Available at: http://www.agi-usa.org/sections/youth.html ; accessed: May 12, 2004.

[5] Card JJ. Teen pregnancy prevention: Do any programs work? Annual Review of Public Health. 1999;20:257-285.

[6] Weinstock H, Berman S, Cates W. Sexually transmitted diseases among American youth: Incidence and prevalence estimates, 2000. Perspectives on Sexual and Reproductive Health. 2004;36(1):6-10.

[7] Chesson HW, Blandford JM, Gift TL, Tao G, Irwin KL. The estimated direct medical cost of sexually transmitted diseases among American youth, 2000. Perspectives on Sexual and Reproductive Health. 2004;36(1):11-19.

[8] Centers for Disease Control and Prevention. HIV Prevention Strategic Plan through 2005. 2001. Available at: http://www.cdc.gov/nchstp/od/hiv_plan/default.htm ; accessed: May 23, 2005.

[9] Advocates for Youth. Serving youth of color. Transitions. Jan 1, 2004; 15(3). Available at: http://www.advocates foryouth.org/ publications/ transitions/transitions1503.htm ; accessed: Sep 8, 2005.

[10] Centers for Disease Control and Prevention. Cases of HIV Infection and AIDS in the United States, 2002. HIV/AIDS Surveillance Report. 2003; 14(2002):1-40. Available at: http://www.cdc.gov/hiv/stats/hasrlink.htm ; accessed: Feb 26, 2004.

[11] Houston Department of Health and Human Services. AIDS: Summary of Houston/Harris County Cases, Cumulative - Reported 1981 through 6/30/2005. Jan 9, 2004. Available at: http://www.houstontx.gov/health/ HIV-STD/ 2ndQRT2005.pdf ; accessed: Sep 8, 2005.

[12] Whalen, L. G., Grunbaum, J. A., Kinchen, S., McManus, T., Shanklin, S. L., and Kann, L. Middle School Youth Risk Behavior Survey 2003. Web. 2005; 1-292. Available at: http://www.cdc.gov/healthyyouth/yrbs/middleschool 2003/ pdf/fullreport.pdf; accessed: Nov 8, 2006.

[13] Kaestle CE, Halpern CT, Miller WC, Ford CA. Young age at first sexual intercourse and sexually transmitted infections in adolescents and young adults. American Journal of Epidemiology. 2005;161(8):774-780.

[14] O'Donnell L, O'Donnell CR, Stueve A. Early sexual initiation and subsequent sex-related risks among urban minority youth: The Reach for Health study. Family Planning Perspectives. 2001;33(6):268-275.

[15] Coker AL, Richter DL, Valois RF, McKeown RE, Garrison CZ, Vincent ML. Correlates and consequences of early initiation of sexual intercourse. Journal of School Health. 1994;64(9):372-377.

[16] Flanigan CM. Sexual activity among girls under age 15: Findings from the National Survey of Family Growth. In: Albert B, Brown S, Flanigan CM, eds. 14 and Younger: The Sexual Behavior of Young Adolescents. Washington, DC: National Campaign to Prevent Teen Pregnancy, 2003:57-64.

[17] Boekeloo BO, Howard DE. Oral sexual experience among young adolescents receiving general health examinations. American Journal of Health Behavior. 2002;26(4):306-314.

[18] Stone N, Hatherall B, Ingham R, McEachran J. Oral sex and condom use among young people in the United Kingdom. Perspectives on Sexual and Reproductive Health. 2006;38(1):6-12.

[19] Halperin DT. Heterosexual anal intercourse: prevalence, cultural factors, and HIV infection and other health risks, Part I. AIDS Patient Care and STDs. 1999;13(12):717-730.

[20] Koniak-Griffin D, Lesser J, Uman G, Nyamathi A. Teen pregnancy, motherhood, and unprotected sexual activity. Research in Nursing and Health. 2003;26(1):4-19.

[21] Manlove J, Franzetta K, McKinney K, Papillo AR, Terry-Humen E. No Time to Waste: Programs to Reduce Teen Pregnancy among Middle School-Aged Youth. Washington, DC: National Campaign to Prevent Teen Pregnancy, 2004.

[22] Borawski EA, Trapl ES, Lovegreen LD, Colabianchi N, Block T. Effectiveness of abstinence-until-marriage sex education among middle school adolescents. American Journal of Health Behavior. 2005;29(5):423-34.

[23] Dilorio C, Resnicow K, McCarty F, De AK, Dudley WN, Wang DT, Denzmore P. Keepin' it REAL! Results of a mother-adolescent HIV prevention program. Nursing Research. 2006;55(1):43-51.22. Thomas B, Mitchell A, Devlin M, Goldsmith C, Singer J, Watters D. Small group sex education at school: The McMaster Teen Program. In: Miller BC, Card JJ, Paikoff RL, Peterson JL, eds. Preventing Adolescent Pregnancy. Newbury Park, CA: Sage Publications, 1992:28-52.

[24] Howard M, McCabe JB. Helping teenagers postpone sexual involvement. Family Planning Perspectives. 1990;22(1):21-26.

[25] Aarons SJ, Jenkins RR, Raine TR, El Khorazaty MN, Woodward KM, Williams RL, Clark MC, Wingrove BK. Postponing sexual intercourse among urban junior high school students-a randomized controlled evaluation. Journal of Adolescent Health. 2000;27(4):236-247.

[26] Coyle KK, Kirby DB, Main BV, Gomez CA, Gregorich SE. Draw the Line/Respect the Line: A randomized trial of a

middle school intervention to reduce sexual risk behaviors. American Journal of Public Health. 2004;94(5):843-851.
[27] Nielsen J. Usability Engineering. San Diego:Morgan Kaufmann, 1993.
[28] Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice Hall, Inc., 1986.
[29] McGuire W. Social psychology. In: Dodwell PC, ed. New Horizons in Psychology. Middlesex, England: Penguin Books, 1972:219-242
[30] Flay BR, Petratis J. The theory of triadic influence: A new theory of health behavior with implications for preventive interventions. Advances in Medical Sociology [4: A reconsideration of models of health behavior change], 19-44. 1994. Greenwich, Conn., JAI Press.
[31] Asker P, Yavuz H, Koksal M. Student's perceptions of computer-assisted instruction environment and their attitudes towards computer-assisted learning. Educ Res. 1992;34:133-9.
[32] Sapperstein MS, Gold RS, Atkinson N, Peterson KJ. BodyFun Online: How usability studies can improve internet-based health education. Am. Pub. Hlth Assoc. 132nd Annual Meeting, Washington DC, Nov. 2004.
[33] Serious Games Initiative. Available at: http://www.seriousgames.org/index2.html. Accessed Dec 3rd, 2006.
[34] Games for Health. Available at: http://www.gamesforhealth.org/index3.html. Accessed Dec 3rd, 2006.
[35] Cook TD and Campbell DT. Quasi-Experimentation. Design & Analysis Issues for Field Settings. 1979. Houghton Mifflin Co. Boston, 1979.

Address for correspondence

Ross Shegog, PhD.
Assistant Professor
Center for Health Promotion & Prevention Research
UT-School of Public Health
7000 Fannin, Suite 2668
Houston, TX 77030
Phone: 713-500-9677
Fax: 713-500-9602
e-mail: Ross.Shegog@uth.tmc.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

HeartCareII: Home Care Support for Patients with Chronic Cardiac Disease

Patricia Flatley Brennan [a], Gail Casper [a], Susan Kossman [a,b], Laura Burke [c]

[a] *University of Wisconsin-Madison, Madison, WI*
[b] *Illinois State University, Normal, IL*
[c] *Aurora HealthCare, Milwaukee, WI*

Abstract

Systematic engagement of patients in disease management requires design and deployment of innovative technologies that complement and extend professional nursing services. We describe here a model of nursing practice that capitalizes on a web-based resource (HeartCareII) to support patient self-management, symptom interpretation, and self-monitoring. Research staff provided computers and technical assistance; visit nurses trained patients in the components of the HeartCareII Website most relevant to their care needs. This paper describes the nursing practice model and the web resource, and reports the experience of patients recruited in the early phase of the study.

Keywords:

heart disease, consumer health informatics, home care services, nursing models

Introduction

Advances in computerized home monitoring could lead to a disease management approach in which the patient is a passive recipient in the assessment and management of key health parameters. While the benefits of passive home monitoring strategies are recognized, the philosophy underlying our approach is that technology should enhance and facilitate patients' *active* engagement in self-monitoring and self-management. Research demonstrates that Web-based information resources that not only alert patients about what symptoms to expect but also provide coaching about how to manage them improve self-care and health outcomes.[1, 2] We propose to create a new model of home care nursing, Technology Enhanced Practice (TEP), and support it with a website, HeartCareII, accessible by *both patients and nurses*. HeartCareII provides tailored coaching information, communication with peers and professionals, and personal monitoring tools designed to make patients more active co-creators of their health and to facilitate effective nursing care. In this study, we integrate the technology within the formal care delivery system, building upon the parallel approach used in our prior work. This paper describes the core of the intervention, and reports early evidence from the field experiment of the ways in which 24 patients who have completed the study used the HeartCareII resources.

Patients with complex cardiac disease, including Congestive Heart Failure (CHF) Acute Myocardial Infarction (AMI), Coronary Artery Bypass Graft (CABG) and Valve Surgery and various combinations of these conditions are the targets of this HeartCareII project. Clinical management of these chronic heart conditions involves a partnership between patients, physicians and nurses. Therapeutic interventions may include pharmacotherapy, diet, and activity consultation[3, 4] and may rely on care provided by home care nurses. Positive effects often result from nursing interventions and home-care approaches, but these interventions are expensive, may over-utilize scarce professional resources, and have effects that cease once the home care nursing interventions are stopped. Even under optimal management, patients with chronic heart disease experience frequent, serious exacerbations. The challenge is to institutionalize an intervention strategy that provides patients with timely access to relevant health information, self-management guidance and self-monitoring resources, while facilitating contact with clinicians for rapid intervention. Consumer health informatics (CHI) solutions, appropriately fit to the care situation and nursing practice model, may improve self-management, reduce demand on hospital and home care services and ensure timely, appropriate treatment.

The quality chasm and nursing shortage require sophisticated information solutions to replace the naïve view that simply providing WWW-based health information will lead to improved disease management and adherence to healthy behaviors. Safran's work demonstrated that the integration of professional presence within an electronic outreach intervention contributed strongly to desirable clinical outcomes.[1] Studies of the equivalence of technology-mediated care with traditional approaches [5] are no longer adequate. They must now be superseded by complex large-scale studies that examine the integration of CHI within existing care approaches, rather than supplanting existing care delivery models with electronic substitutes. Recent evidence suggests that tailoring the information provided in the technology intervention to the patient's health status and learning needs improves outcomes. [6, 7] In this study we draw information regarding the needs for tailored health information from the patient *in addition to* considering how information technology resources complement or expand resources available in the nurse-patient encounter. In creating TEP we provide a

Selected for best paper award.

means to close a gap among consumer health informatics tools, evidence-based nursing practice, and home health nursing practice by integrating the informatics innovation into the routine process of home care nursing. [8] Practice innovations such as TEP may provide sufficient support to ensure that all nurses have the competence to address the needs of the person with heart disease.

Materials and methods

HeartCareII project

The goal of the HeartCareII project, a collaboration between an academic nursing informatics research team, a health system clinical research group and a practice setting, is to develop and test a sustainable model of home care nursing supported by a technology core and home based technology. The practice setting is a large metropolitan visiting nurse association (VNA). We believe home based technology will help the nurse provide better care during home visits and also support extended asynchronous care.

The innovative home care nursing practice model labeled, *Technology Enhanced Practice*, provides an approach that allows home care nurses to balance interpersonal and technologically-mediated interventions in the care of patients with heart disease, thus maximizing their time and effectiveness while supporting patient's self-management efforts.[8] The major aim of the HeartCareII project is to determine whether TEP leads to improved outcomes in patients with heart disease. This multiyear project, conducted within a large integrated health care delivery system in Wisconsin, will involve 300 patients and approximately 80 home care nurses in the study's five VNA offices. Most nurses are female; the average age is 43 years old (range 24-60). Racial composition is predominantly Caucasian. Although all nurses use computers for planning and documenting the care they provide, they exhibit considerable variation in their level of comfort and skill in using them.

The first phase of the study, the design phase, applied human factors techniques to analyze the work of home care nurses. It resulted in development of the technology core for TEP, the HeartCareII website, which is housed within the clinical partner's patient portal. The second phase of the study, initiated in August 2005, implements the TEP and its technical core, the HeartCareII website, in a randomized field experiment that will examine its impact on nurses' workflow and patient outcomes including clinical status, quality of life, self-monitoring, satisfaction with care, and health care service utilization.[9]

Nursing practice models

We envision TEP as a new nursing practice model for home care nurses. Nursing practice models (NPMs) can be broadly defined as the "manner in which nurses assemble to accomplish clinical goals".[10] They represent structural and contextual features of care settings, including organizational, nursing resource and support dimensions[10,11] No universal, ideal NPM exists and there are many variations across care settings. Additionally, actual practice may deviate from anticipated practice specified in organizational NPM plans. NPMs share common elements that provide a framework to understand factors that may support, or interfere with, nursing practice. Basic elements common to NPMs and some associated attributes include 1) nurses' role (autonomy, accountability, philosophy of care); 2) professional relationships (teamwork, collaboration, communication, coordination); 3) care delivery systems (primary, case management); 4) continuity of care; 5) management systems (shared governance, decentralized decision making); and 6) compensation.[10-13] The degree to which the attributes listed in each category are present influences nurses' ability to administer effective interventions. Ideally, nurses have the autonomy and accountability to make appropriate care decisions, effectively communicate and collaborate with healthcare team members and patients, ensure continuity of patient care through adequate staffing with skilled and knowledgeable nurses, and be active participants in decision making.[13] In the HeartCareII study, the homecare nurses' existing NPM served as the framework on which to build the technology nurses could use in providing TEP. We designated this existing NPM as the "usual care" practice for the purposes of the experiment.

Usual care

The usual care for patients with heart disease is provided by Registered Nurses accountable for delivering evidence-based nursing care to homebound patients based on their assessment of the patient's needs, physician's orders and clinical guidelines developed by advanced practice nurses. The patient's diagnosis, acuity, physician's orders, insurance coverage and nursing care needs determine the number of home visits to be made, which typically ranges from 1-9. Continuity of care is provided by having a single nurse provide as many of the home visits as possible.

Nurses interact with patients through home visits and telephone for the specified number of visits typically spanning several weeks. The usual nursing care for patients with any diagnosis of chronic heart disease includes initial and ongoing comprehensive health assessments, patient and family education, medication management and administration, skilled interventions, and coordination of other necessary health services. Patient and family education are integral components of care. Nurses use an institutional set of care management initiative education materials to assist patients to adhere to medications, modify their life style, understand the disease, and recognize early manifestations of disease progression or complications. In addition to education, the practices include surveillance of cardiopulmonary symptoms, and prompt response to a change in status. Nurses coordinate services which include dietary, intravenous therapy, laboratory, physical and occupational therapy, behavioral health, referral to the VNA heart failure program, cardiac rehabilitation services, and telemanagement.

Nurses at this VNA use a set of guidelines as the benchmark for usual care. These guidelines are based on Barrella and Monica's Congestive Heart Failure at home clinical pathway [14] and are extensible to patients with many

types of chronic cardiac disease. The clinical pathway aids the nurse in developing a plan of care to facilitate patient integration of, and compliance with, the medical plan of care. Practice innovations in combination with such guidelines may provide sufficient support to enable nurses to have the competence required to address a full range of patient needs.

Nurses are organized in care teams of one to three nurses and a care coordinator; five VNA offices located in southeastern Wisconsin participate in the study. A single nurse takes primary responsibility for a given patient's care plan and makes as many of the visits to that patient as possible. Communication and collaboration occur during weekly nursing care team meetings and monthly staff meetings. Nurses also collaborate with physicians and other healthcare team members as needed. The VNA follows a shared governance structure, and nurses are paid per visit incorporating visit intensity.

Technology enhanced practice

Technology Enhanced Practice is a nursing practice model in which nurses selectively and deliberately employ computer and information technologies in a manner designed to meet individual patient care goals. We built the HeartCareII website to serve as the technology core for TEP based on the existing nursing practice model and added resources to expand, support and supplement nursing work, to enhance nursing care and to facilitate patient self-monitoring and self-management. Existing paper based tools nurses use to provide usual care were modified to an electronic format and supplemented with additional tools to address care needs (e.g. a drug checker, several health trackers).

The technology core (HeartCareII website) provides a standard suite of technology services which nurses can select to use based on patient needs and illness trajectory.[9] Additionally, we anticipate that the technology features that support TEP may affect NPM elements such as nurses' role, communication, collaboration and continuity of care. The HeartCareII website is housed within the patient portal section of the clinical partner's public website and may be accessed by nurses and patients through an internet connection from a variety of geographic settings (point of care, physician or VNA office, remote locations such as cardiac rehabilitation centers or family members' homes). Interactive tools on the HeartCareII website address symptom monitoring (symptom checklist, weight tracker, blood pressure tracker, heart rate tracker), education (patient education papers on a variety of topics, and trackers such as the food and fluid trackers), and communication (my goals, my journal, email, bulletin boards). A sample display from the HeartCareII website is provided in Figure 1.

Nurses encourage patients to use pertinent tools daily and to explore the website on their own. The nurse can monitor the patient remotely by accessing their account (with permission) to check on symptom monitoring, and communicate in person, on the telephone, or by computer. Additionally, patients may communicate with other study participants through a private bulletin board.

Nurses' actual use of the technology resources in the delivery of patient care (e.g. type of service, frequency, and duration) varies both within an individual nurse's caseload and among different nurses based on patients' needs and abilities as well as nurses' levels of clinical expertise, preferences, and experience with technology. Nurses use professional judgment to individualize TEP for each patient based on their skills, preferences and assessment of patient needs. The following description illustrates how a nurse might use the TEP intervention during the home visit phase:

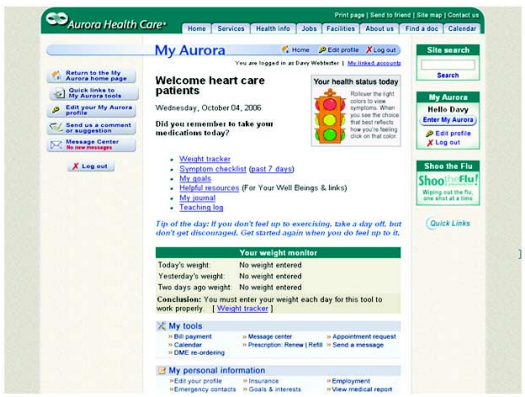


Figure 1 – The HeartCareII Home Page

Upon assignment of an eligible patient to his or her caseload, the nurse will complete the standard VNA intake protocol and conduct an information needs assessment. Nurses will use this information to plan a TEP intervention for this patient, identifying practice interventions and selecting features of the computer support for the patient. If the nurse appraises this patient information needs as high, he or she may use the HeartCareII website to provide ongoing, between-visit education for the patient, directing the patient to relevant readings. If he or she also determines this person has strong needs for coaching and support, and he or she may use the HeartCareII web site to message the patient daily, alternating home visits with daily electronic monitoring and communication, thus ensuring that the patient has adequate support. Thus, TEP supports creative use of human resources and information technology to meet patient needs.

Conduct of the HeartCareII experiment

Through a field experiment we will compare selected outcomes of TEP to those of usual home nursing care. Three of the VNA offices are designated as "experimental sites" and deliver care to patients with chronic cardiac disease using TEP; the remaining two VNA offices provide usual care. Patients in the TEP arm receive computers and related software or have their own computer reconfigured for the study; patients in the usual care condition receive a booklet including clinical care guidelines, standard patient teaching materials and paper forms for documenting health parameters. The study period for each patient lasts six months, with a home care phase during which the nurse delivers TEP (about 2 to 8 weeks) and the post-discharge phase in which the patient assumes responsibility for self-

monitoring and self-management that lasts the remainder of the six month study period.

Results

Sample

In the first 15 months of the experimental phase of the project, we enrolled 152 patients and 47 have completed the 6-month study period (24 in the experimental group and 23 in the comparison group). We report here on the 24 patients in the experimental group who have completed the study to date. Patients included in this analysis range in age from 43 to 88 years old; mean age is 69 (S.D. 11.6). There are 10 women and 14 men. Fifteen patients have a primary diagnosis of CHF, with the other nine patients recovering from coronary artery bypass graft or valve replacement surgery.

We used the Specific Activity Scale (SAS) [15] to measure functional capacity. Fifteen respondents reported being able to walk down a flight of steps without stopping; 17 were able to dress themselves without stopping because of symptoms. Only three of the patients were able to strip and make a bed. The remaining items (ability to walk up steps carrying something or carrying 25 pounds) of the SAS require careful interpretation because inability to do these activities may be related medical restrictions rather than functional capacity.

Patient diagnoses included congestive heart failure (CHF), coronary artery bypass graft surgery (CABG), aortic valve replacement and heart transplant. Their experience with computers prior to participation in the study ranged from "none" to "competent". Length of time in the study while receiving home nursing care ranged from one week to greater than two months. Approximately one-third of the patients used the Internet access device and dial-up connection provided by the study (Wyse thin client with a CRT); the remaining two-thirds used DSL connections with the study device or with their own computers.

All patients in the experimental arm received TEP; however, the extent to which they used the HeartCareII website varied, and examining this use in depth will provide a foundation to later characterize TEP. Thus, we turn to an in depth characterization of use of the HeartCareII website. Because the study is on-going, we will not be providing any outcome data in this report.

Patient use of the HeartCareII website

Patient use data are collected through passive capture of logins and webpage tracking. Login counts are reported weekly, while webpage tracking is provided "on-demand". Accuracy is validated by comparing the two sources of data.

Use of the HeartCareII resources is highly variable. Logins to the system (defined as a successful connection to the site after submitting user ID and password) range from two to 144 with a mean of 36. Length of use of the system (days between first and last login) varies from zero to 194 days, with a mean of 86 days. Two-thirds of the patients used the HeartCareII website for 50 days or less and tended to have 20 or fewer logins. Approximately one-third of the patients used the system for greater than 50 days and tended to have more logins. Once connected to the HeartCareII website, the most commonly used resources included the weight tracker, food tracker and an action plan that assisted in symptom interpretation. Figure 2 below plots the number of logins against the number of days the patient used the system.

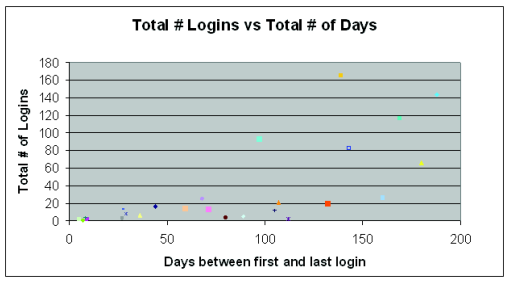


Figure 2 – Plot of Use Interval with Login Count

Illustrative case studies

The following case studies summarize two patients' actual use of the HeartCareII website resources and illustrate nurses' and patients' use of the website during the home visit and self-management phases. These cases provide beginning insights into the nature of TEP.

Case 1: Mr. A. was admitted to home care services nine days after receiving a heart transplant. He consented to participate in the HeartCareII study and had the study-provided internet device installed 11 days later. He had used a computer for personal and professional purposes and described himself as "familiar" with using computer technology.

Although there was no nursing documentation of his instruction in the use of the HeartCareII website or evidence of collaborative use by any of his three nurses, Mr. A. used many of the self-monitoring and information tools extensively, logging in almost daily for periods of 35 to 110 minutes. He even reported returning the study-provided computer and dial-up ISP after purchasing a PC with cable Internet to make his access easier and faster. He primarily used the health trackers - recording his weight, fluid and sodium intake, and blood pressure daily. He also reported using the HeartCareII website to learn about his diet and to seek information about heart disease. He continued to access the HeartCareII website for two months after his discharge from home care.

Case 2: Following his hospitalization for heart surgery, Mr. F. was referred to the VNA for post-operative care and instruction. He was enrolled in the study four days later. A PC owner with cable Internet connection, Mr. F. described himself as "competent" in the use of a computer. He opted to use his own computer for access to the HeartCareII website.

Due to limited insurance authorization, Mr. F.'s nurse was able to make only one home visit. Following that visit, she documented that he was able to access the website, locate information, and that he "found the information valuable". Mr. F. logged on to the HeartCareII website only two times prior to his discharge from the VNA.

Discussion

We envisioned a typical course of TEP for an experimental patient in the HeartCareII study to last for six months, encompassing about one week to two months of intensive nurse-patient interaction (laying the groundwork for subsequent self-care management by the patient and his or her family caregiver) and approximately four months of patient self-management aided by the HeartCareII website. We envisioned that use of the resources would vary over time and by individual. Our observations to date confirm that duration of VNA service and use of the HeartCareII resources does indeed vary across patients and nurses. Patients' use of the HeartCareII tools reflect that about two-thirds were engaged with the technology for 50 days or less, while one-third continued to use the resources for the entire 6 month period that it was available to them. About 25 percent of these opted to keep the internet device provided by the study, suggesting potential continued use of the electronic resources beyond the study period.

The variability in use of the HeartCareII resource follows patterns of use that parallel those found in earlier studies.[16,17] They provide evidence supporting patients' active engagement in web-based information and use beyond the initial home visit phase, suggesting that this informatics tool extends nursing care impact.

The variability in use of the HeartCareII website may stem from many factors. Patients may have found the technology off-putting and difficult to manage, although this has not been our experience in earlier studies with similar-aged participants.[2] Nurses may vary in their ability to understand and encourage patients to use the HeartCareII website. Or, most interestingly, we may be seeing the beginnings of differentiated practice under a TEP model, in which the use of the technology varies with the nursing goals.

- Nurses' role in encouraging and assisting patients to use the HeartCareII resources remains an important, yet unexplored explanatory aspect of use. Follow-up work system analyses are planned to further evaluate delivery of TEP.
- Evidence presented here demonstrates that home care patients with chronic cardiac disease can and will use a web-based resource for self-management, self-monitoring and clinical communication. Explication of the full nature of technology enhanced practice and determining its impact on patient outcomes awaits completion of this study.

Acknowledgments

Support for this project was provided by the National Library of Medicine LM 6247 and Aurora Health Care.

References

[1] Gray JE, Safran C, Davis RB, Pompilio-Weitzner G, Stewart JE, Zaccagnini L, et al. Baby CareLink: using the internet and telemedicine to improve care for high-risk infants. *Pediatrics*. 2000 Dec;106(6):1318-24.

[2] Brennan PF, Moore SM, Bjornsdottir G, Jones J, Visovsky C, Rogers M. HeartCare: an Internet-based information and support system for patient home recovery after coronary artery bypass graft (CABG) surgery. *J Adv Nurs*. 2001 Sep;35(5):699-708.

[3] Rich MW. Non-pharmacological treatment of heart failure: Just as important? In: McMurray JJV, Cleland JEF, editors. Heart failure in clinical practice. London Martin Dunitz; 1999.

[4] Gheorghiade M, Cody RJ, Francis GS, McKenna WJ, Young JB, Bonow RO. Current medical therapy for advanced heart failure. *Heart Lung*. 2000 Jan-Feb;29(1): 16-32.

[5] Jenkins RL, McSweeney M. Assessing elderly patients with congestive heart failure via in-home interactive telecommunication. *J Gerontol Nurs.* 2001 Jan;27(1):21-7.

[6] Strecher VJ. Computer-tailored smoking cessation materials: a review and discussion. *Patient Educ Couns*. 1999 Feb;36(2):107-17.

[7] Revere D, Dunbar PJ. Review of computer-generated outpatient health behavior interventions: clinical encounters "in absentia". *J Am Med Inform* Assoc. 2001 Jan-Feb;8(1):62-79.

[8] Brennan P, Burke L, Casper G, Sebern M, Krause C, Kossman S, et al. Creating technology-enhanced practice: A university-home care-corporate alliance. In: Park H-A, Murray P, Delaney C, editors. Consumer-centered computer supported care for healthy people Proceedings of NI 2006; 2006; Seoul. S. Korea: IOS Press: Amsterdam. ; 2006. p. 664-48.

[9] Casper GR, Karsh BT, Or CK, Carayon P, Grenier AS, Brennan PF. Designing a technology enhanced practice for home nursing care of patients with congestive heart failure. AMIA Annu Symp Proc; 2005; 2005. p. 116-20.

[10] Brennan PF, Anthony MK. Measuring Nursing Practice Models using Multi-Attribute Utility theory. *Res Nurs Health*. 2000 Oct;23(5):372-82.

[11] Mark BA. Characteristics of nursing practice models. *J Nurs Adm*. 1992 Nov;22(11):57-63.

[12] Arford PH, Zone-Smith L. Organizational commitment to professional practice models. *J Nurs Adm*. 2005 Oct;35(10):467-72.

[13] Hoffart N, Woods CQ. Elements of a nursing professional practice model. *J Prof Nurs*. 1996 Nov-Dec;12(6):354-64.

[14] Barrella P, Della Monica E. Managing congestive heart failure at home. *AACN Clin Issues*. 1998 Aug;9(3):377-88.

[15] Goldman L, Hashimoto B, Cook EF, Loscalzo A. Comparative reproducibility and validity of systems for assessing cardiovascular functional class: advantages of a new specific activity scale. *Circulation*. 1981 Dec;64(6):1227-34.

[16] Brennan PF, Moore SM, Smyth KA. The effects of a special computer network on caregivers of persons with Alzheimer's disease. *Nurs Res*. 1995 May-Jun;44(3): 166-72.

[17] Shaw BR, Han JY, Baker T, Witherly J, Hawkins RP, McTavish F, et al. How women with breast cancer learn using interactive cancer communication systems. *Health Educ Res*. 2006 Jul 7.

Address for correspondence

Patricia Flatley Brennan, RN, PhD, FAAN
University of Wisconsin-Madison
School of Nursing H6/241 CSC
600 Highland Ave
Madison, WI 53792

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

A Web-Based Communities of Practice Support System for Caregivers

Shirley L. Fenton[a,b], H. Dominic Covvey[a], Douglas W. Mulholland[b], Donald D. Cowan[a,b], Judith Shamian [c], Bonnie Schroeder[c]

[a] *Waterloo Institute for Health Informatics Research, University of Waterloo, Canada*
[b] *Computer Systems Group, University of Waterloo, Canada*
[c] *Victorian Order of Nurses (VON), Canada*

Abstract

CareNet is an interactive Web-based system intended to support informal caregivers (ordinary citizens who are engaged in providing residential healthcare to their families and friends). The design of CareNet uses concepts from a number of areas including: Communities of Practice, software engineering, content authoring, and knowledge management to create a supportive environment for the caregiver. The specific objectives of the CareNet project are: (1) To create a highly effective interactive environment that addresses the needs of caregivers to: (1a) obtain information and guidance, (1b) achieve efficient communication with professional healthcare workers with whom they collaborate and with other caregivers from whom they can gain advice, emotional support and solace, (1c) access other information and physical resources that are needed for proper care, and (1d) document their observations, interventions and insights that can in future become a knowledge resource for other caregivers; and (2) To demonstrate the beneficial impact of CareNet on caregivers who collaborate with professional care providers and on their interactions with professional care providers.

Keywords:

Communities of practice (CoP), caregivers, web-based system, consumer e-health, educational systems, Internet, Web, knowledge management.

Introduction

Healthcare faces many challenges if it is to remain affordable and accessible. Specifically, we require innovative approaches and alternatives to institutional solutions if we are to keep costs under control. One such solution is to discharge patients with short-term illnesses as well as individuals with chronic conditions earlier than in the past with the home, family and friends then becoming the locus for recovery. Using the home as the centre of care is definitely a trend as patients are discharged from institutional care at a higher level of acuity than the levels at which they were admitted just a few decades ago. This approach imposes many challenges on these "informal caregivers," who generally lack adequate care-related knowledge and skills, and are already living a full and often independent existence.

How can we provide adequate support to these informal caregivers? Can modern information and communications technology, such as the web, provide useful tools and valuable experiences? In this paper, we present a description of such an interactive web-based information systems called "CareNet." In addition to describing the technology, we will explain the motivation for and the process of arriving at the CareNet solution and the preliminary results that have already been observed. This project has motivated us to look more deeply into web-based information systems that can support the informal caregiver, as well as other communities of practice, that exist in healthcare and elsewhere. Our current work continues to examine the information technologies that would be useful in supporting communities of practice and how members of such communities interact with these technologies.

The Waterloo Institute for Health Informatics Research (WIHIR) and the Computer Systems Group, both at the University of Waterloo (UW), worked closely with the Victorian Order of Nurses (VON) to develop a web-based system to support informal caregivers in:

- providing effective and sustainable care in the home setting and
- achieving effective and efficient communications with professional healthcare personnel when needed.

CareNet is intended to provide informal caregivers with access to "support channels," such as care-related information resources, and the means to communicate with professional care providers and other informal caregivers in similar situations. CareNet is intended to provide experience-based insights of caregivers and providers, directories of physical resources (e.g., pharmacies), event co-ordination via booking and scheduling services, and other capabilities. All of these aim to reduce the burden of care giving and increase the interaction with other caregivers who can provide assistance and emotional support.

Our approach allowed for a team of professional caregivers from the Victorian Order of Nurses (VON) to be completely involved in the requirements, design and development of the system using a rapid prototyping tool developed at the University of Waterloo called the Web-

based Informatics Development Environment (WIDE)[4]. The resulting system (CareNet) is an example of a Community of Practice support system[7],[8],[9],[10],[11],[12] which facilitates interaction among caregivers, and provides information and other resources. According to Wenger et al., "Communities of practice are groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly."

Our research objectives were to:

a) Determine and document caregivers' detailed needs;
b) Define the capabilities required of CareNet to address the needs of caregivers and VON Canada;
c) Design and implement a proof of concept version of CareNet;
d) Apply CareNet to address key needs of patients and providers; and
e) Evaluate the impact of CareNet on caregivers and care professionals.

Methods

Our approach to developing CareNet included the following steps:

1. Form a small joint UW-VON Project Team to develop the prototype. The VON members of the team were individuals who were knowledgeable about "informal caregivers" and the type of issues they face.
2. Identify the goals and objectives that VON staff believes address the needs of informal caregivers who currently seek advice from and collaborate with VON professionals in care provision.
3. Analyze these goals and objectives and use the analysis to define the needs of caregivers that can be supported by a web-based interactive information system.
4. Design the system and review this design with the VON. This step could go through several iterations as both the VON staff and the systems designers develop a more comprehensive understanding of the issues related to the problem.
5. Create a pilot version of the CareNet portal suitable for demonstration to caregivers and VON professionals using a set of rapid development frameworks developed at the University of Waterloo called the Web-based Informatics Development Environment (WIDE). Web-based information systems created using WIDE can be built and modified quickly. Thus the system itself can be and was used as a requirement elicitation tool to refine the design further with the feedback received from the iterations. Shared elicitation and design sessions with VON staff and development staff were held where the two groups were remote from each other. The nature of WIDE also allowed system changes on the fly thus supporting "what-if" scenarios.
6. Review the pilot version by VON staff supporting caregivers and informal caregivers. This review would help determine whether the functionality of the system meets user needs.

Once the goals and objectives were identified and analyzed, a number of applications for the prototype were discussed and prioritized. From this analysis and prioritization, it was determined that three primary applications would be created for the prototype:

- A directory of services offered by VON branches across Canada.
- Information resources for the caregiver community.
- A tool to allow both caregivers and supporting professionals to augment the online content based on personal experiences and to make a vetted version of this content available to both parties.

A description of each of these applications will be presented later in this paper.

As mentioned earlier in this paper the system design and implementation was developed iteratively and interactively by the UW and VON Project Team members. Each step of the implementation was reviewed with VON staff responsible for the project and with some staff members who are care professionals. One of the design criteria specifically focused on maintenance of the information handled by the CareNet portal. It was intended that all information available through the portal could be maintained and manipulated by VON staff involved in healthcare and that the training required would be no more complex than learning to use the basic functions of a word processor. This included vetting and manipulating all the online content that was supplied by the users including both informal caregivers and supportive VON staff.

The project was not field tested with informal caregivers as there was not enough information available in the prototype system to support an adequate test. However, it was felt that the variety of experiences that the VON staff had with informal caregivers would provide a sufficient level of validation at this moment in time, and with the rapid development nature of the WIDE platform, essential changes to CareNet could be identified later and easily implemented. Based on the results of this research project, the VON management has elected to contract the development and operation of a complete system to a Canadian company.

Results

The CareNet portal is more than just an information source and chat room. It provided a test bed for a number of concepts related to user support and also to methods of software development in domains where the experts themselves are still trying to understand the problem at hand. From the user perspective, the CareNet portal is a proof of concept of:

1. a Community of Practice support system. The CareNet portal supports two identifiable groups of people who share a concern or need for the delivery of informal care namely the informal caregiver and a specific distributed group of VON staff members. They come from different perspectives, bring different expertise to the community, and can be mutually supportive of each

other. The system itself allows interaction on an as needed basis, providing reinforcement, assistance and interaction on demand.

2. an experimental platform for supporting caregivers. The platform allowed the members of the community of practice both to share existing information and also to add new sources and experiences. It is recognized that new information would need to be vetted and automatically indexed to ensure that inappropriate information is not inserted into the system and that it can be easily found at a later date. However, it should be possible through appropriate controls to ensure that vetting of information would not be an onerous task.
3. a web-based system which can be maintained and enhanced by VON staff with minimal technical expertise support. This property is key to the successful operation of any web-based information system, particularly the interactive one that we have proposed. The information experts should have the ability to maintain the content themselves with little or no help from technical support staff. This property is essential to the ongoing maintainability of the system.
4. a rapid development environment for web-based applications. With the advent of more public systems such as the web, we are exploring many web applications with wider audiences. In many cases, these applications are not well understood and need to be modified frequently based on the experience and feedback from both the users and the service providers. Thus a rapid development environment is an essential tool where the information system can be modified quickly and reliably, preferably while it is in operation.

Application details

Three applications were developed for the CareNet portal prototype which is now operational. The applications are quite advanced and include the following capabilities:

VON directory

This application created a searchable directory of VON branches and their services. The information on each branch includes a list of its services as well as a map showing its street location. The directory can be viewed in English and in French. The directory database is maintained online. Maintenance of the directory information can be done easily by authorized central or branch staff using simple web-based forms. This directory would not only be useful for the general public seeking the closest VON branch or the branches which offer a particular service, but it is also an online resource for VON staff in answering client questions and a tool to analyze VON resources. For example, “How many branches offer certain services?;” “Where are possible gaps in services?;” “Should names of services be consistent?;” and if so “What should they be called?”

Caregiver support system

To create this application, three types of information a caregiver needs were considered:

a) Information about a particular disease or health-related problem (Constraints).
b) What organizations can provide help (Resources).
c) How to do something (Protocols).

Collectively, the Constraints, Resources and Protocols were called CPR. A system was designed to include these three types of information resources. This application was based on a section of the "Resource Guide for Family Caregivers", a manual [5],[6] produced by individuals associated with the Family Caregivers Network Society in Victoria BC. A prototype interface was designed to determine the caregiver’s request and display the appropriate information resource or set of resources. The system can provide information that provides support for both the caregiver and the recipient of the care. For example, it handles questions such as: “How do I get support so I can take a break from my care giving role?,” (Resource) or “What kind of help can I get for my parent with Parkinsons disease? ” (Constraint, Protocol)

Caregiver information and knowledge sharing tool

There is a wealth of useful information that professionals have not documented or that caregivers have learned from personal experiences. This application allows both caregivers and professionals to document this knowledge and information. The tool was designed to allow the collection of this information and to make it available so that others could share and gain value from this information in a simple way. The initial tool is based on a discussion forum. The information is categorized (indexed) and can be searched. Such a tool normally would also contain a vetting mechanism to assist in determining appropriateness with perhaps some editing capability.

The WIDE toolkit

The CareNet system was developed using the WIDE Toolkit, a set of technologies that have been used to construct frameworks[3] for over 30 portals. The WIDE Toolkit contains the following functionality:

- The ability to connect to multiple distributed, disparate databases. This function allows each VON site to produce and maintain databases for local services while supporting a global view of the information.
- An indexing and search engine This function supports indexing and subsequent searching of all databases and documents placed in the portal.
- Structured database content presentation (listings, reports). Reports such as a listing of services and where they are available can be easily introduced and modified.
- Database content administration (remote updates). Data can be changed by any authorized person from any location.
- Structure administration (maintenance of content presentation). The entire WIDE toolkit is built on the premise that we may want to modify both the presentation and the structure of the application.

- Online distance education tool. Succinct online documentation is a key to portal usability and maintenance. The system incorporates an easy-to-use online educational tool that provides the user with the ability to read and annotate documentation.
- Geographic mapping of data. The mapping tools are simple to implement and use but are quite powerful. They allow viewing of data by geographic location, searching of an area for geographic data with specific properties and posting and recording of geographic data. The mapping tools are not based on geographic information systems (GIS) but can communicate with them in both directions (read, write) through accepted standards.
- Interactive visual display (display charts and diagrams constructed from database content). The display tools are derived form the mapping tools and have many of the same properties.
- Access control framework. Access to the data and various reports and other displays is role based. Thus a person can easily be included and excluded from changing data.
- Notifications of events. Based on criteria, users can request notification by e-mail of additions, changes or deletions to content of interest.
- Agents. Agents typically play housekeeping roles and can perform functions such as reporting and/or repairing discrepancies between databases/web sites automatically.
- A self-assessment framework. The self-assessment framework which is an application of WIDE can be populated with data appropriate to a particular field of study and the user can determine their knowledge in that field. The framework also connects to supplementary materials for self-study.

Discussion

CareNet provides caregiver support and demonstrates what can be gained by using tools, such as the WIDE Toolkit, that reduce the barriers to end-user development and sustainability[1],[2]. This is particularly important when new types of requirements and little user knowledge of potential solutions exist. CareNet allows significant systems support in the home setting using browser accessible technologies.

This research project is a stepping stone to further investigation. Many questions remain unanswered and new issues became evident that should be explored. For example, what are the best methods for the capture of undocumented organization information? How can we easily classify this type of information to enable the rich retrieval of this information? What are effective mechanisms to capture the needs of the caregiver? How do we enable the professional caregiver organizations to use this system as a powerful analytical tool?

Conclusion

Home and community care organizations can benefit from applications systems like CareNet and toolkits that allow affordable and rapid system construction and revision. If components of the EHR are to be sourced from the community setting, solutions like these are essential.

Finally tools such as the web provide new methods of supporting communities of practice. What are the most effective way of using these tools? Many unanswered questions still remain to be explored in this area as well.

Acknowledgments

This work was supported by the VON Canada.

References

[1] Cowan DD, Fenton SL, Mulholland DW. Lowering technological barriers to developing, deploying and maintaining web-based interactive information systems. International Journal of Technology, Knowledge and Society 2005; 1(4):175-82.

[2] Cowan DD, Covvey HD, Fenton S, Miller J, Mulholland D, de Groot M, Matyas C, McGarry F. Sustainability models for web-based information systems supporting communities of practice. International Journal of Technology, Knowledge and Society, to appear.

[3] Cowan DD, Alencar PSC, Luo M. A framework for community information systems. Annals of Software Engineering (ASE), Special Issue on Object-oriented Web-based Software Engineering 2002; 13:381-411.

[4] Cowan DD, Fenton S, Mulholland D. The web-based informatics development environment (WIDE), http://csg.uwaterloo.ca/wide.htm, 2006.

[5] Family Caregivers Network Society, Resource Guide for Family Caregivers, Victoria, BC, 2006.

[6] Family Caregivers Network Society, Educational Activities to Support Family Caregivers: Facilitator's Manual, Victoria, BC, 2006.

[7] Hildreth P, Kimble C. Knowledge Networks: Innovation through Communities of Practice, Idea Group Publishing.

[8] Kimble C, Hildreth P, Wright P. Communities of Practice: Going Virtual, Chapter 13 in Y. Malhotra (Ed.) Knowledge Management and Business Model Innovation, Idea Group Publishing, Hershey (USA)/London (UK), 2001, pp 220-34.

[9] Kimble C., Li F, Barlow A. Effective Virtual Teams through Communities of Practice, University of Strathclyde Management Science Research Paper No. 00/9, 2000.

[10] Pan Shan L, Leidner D. Bridging communities of practice with information technology in pursuit of global knowledge sharing. The Journal of Strategic Information Systems 2003;12(1): 71-88.

[11] Wenger E, McDermott R, Snyder W. Cultivating communities of practice: A guide to managing knowledge, Harvard Business School Press, 2002.

[12] Wenger E. Communities of practice: Learning, meaning, and identity. Cambridge University Press, New York, 1998.

Address for correspondence

Shirley L. Fenton
Managing Director
Waterloo Institute for Health Informatics Research
University of Waterloo
Waterloo, Ontario, CANADA
N2L 3G1

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Core Features of a Parent-controlled Pediatric Medical Home Record

Roberto A. Rocha[a,b], Alfred N. Romeo[c,d], Chuck Norlin[c,e]

[a]RemedyMD, Inc., Utah, USA; [b]Department of Biomedical Informatics, University of Utah School of Medicine, USA; [c]Utah Collaborative Medical Home Project, Utah, USA; [d]Utah Department of Health, USA; [e] Department of Pediatrics, University of Utah School of Medicine, USA

Abstract

We describe a coordinated effort to identify the core features of a parent-controlled personal health record for children with special health care needs, involving parents, care givers, and healthcare providers. A summary of the core features is presented, emphasizing needs that are not commonly recognized as functions of a generic personal health record. Our goal was to identify requirements for personal records that empower parents to effectively obtain, organize, understand, and communicate the information necessary to help their children receive the best possible care.

Keywords:

Personal Health Record, pediatrics, children with special health care needs, consumer informatics

Introduction

Children with developmental disabilities and other complex medical problems (collectively, "*children with special health care needs*" or CSHCN) comprise about 12.8% [1] of children in the US and are defined as "those who have or are at increased risk for a chronic physical, developmental, behavioral, or emotional condition and who also require health and related services of a type or amount beyond that required by children generally" [2].

Some CSHCN have relatively common disorders, such as asthma and attention deficit hyperactivity disorder, but the majority have uncommon diagnoses and many have associated disabilities. Roughly 6.5% of US children experience some disability as a result of a chronic condition, the most common ones being respiratory and mental impairments [3]. Rates of disability have increased over the past two decades, as have racial disparities in disability prevalence. The rate of activity-limiting disability for white children increased from 4.07% to 5.97% between 1979 and 2000, while the rate for black children increased from 3.79% to 6.71% [4].

The value of comprehensive and coordinated medical care for children with chronic conditions and disabilities has been demonstrated by several studies [5-8]. The importance of such care for CSHCN has led the American Academy of Pediatrics and the federal Maternal and Child Health Bureau to promote the "Medical Home" concept [9]. Ideally, care provided in the Medical Home is "accessible, family-centered, continuous, comprehensive, coordinated, compassionate, and culturally-effective" [10]. The Medical Home was conceptualized around one of the key elements of comprehensive care: ready access to all relevant information about a patient with a chronic or complex medical condition [11].

Typically, care for CSHCN is provided by numerous individuals (physicians, therapists, dentists, counselors, etc.) and institutions (clinics, outpatient centers, hospitals, etc.), and in multiple environments (home, day care, school, etc.). It is rare that these providers and settings share information systems or communicate information in ways that result in comprehensive, coordinated care. The primary care provider ("Medical Home") is envisioned as the "central repository" for collecting, managing, and appropriately sharing this information, but this goal has yet to be realized.

Personal child health records have been promoted by national and international institutions and have been widely used and appreciated by parents for many years [12]. A national survey published in 2006 found that only 21.3% of primary care pediatricians used an "Electronic Health Record" (EHR) [13]. Until there is widespread adoption and use of interoperable electronic records by primary care providers and other providers, schools, etc., we probably should not expect physicians to offer a viable EHR solution to CSHCN.

Parents of CSHCN, particularly those whose children have developmental, physical, or mental disabilities, spend much of their time caring for their child, managing their multiple health care providers and the information generated by them, and interfacing between those providers, insurance companies, day care and schools, interested family members, and many others. The amount and complexity of information that these parents, and the children themselves as they near adulthood, need to understand, remember, and communicate is daunting. Accurate and complete information is critical to assuring the quality and safety of the health care such children receive, optimizing the coordination of care among providers, and minimizing the duplication of diagnostic and therapeutic interventions. Ready access to records of care, instructions, and reliable information about their child's condition and recommended therapies may encourage parental compliance, as well as educate family members or other caregivers. The ability to access, record, and share information with providers can save time, improve care, and prevent errors. In essence, a parent-controlled health record may offer a bet-

ter and more feasible solution that could take advantage of and enhance collaboration with the Medical Home.

Materials and methods

The overarching needs that guided the identification of core features for a parent-controlled "Pediatric Medical Home Record" (*PedMHR*) were derived from the "Personal Health Record" (PHR) principles recently published [14], including: a) parents are ultimately responsible for decisions about their children's health; b) parents should have access to a reliable and complete record of their children's health information; c) parents should have control and accountability over how these records are used and shared; and d) information in the PHR should be understandable to parents and other caregivers.

In order to identify the core features of an electronic record that would be perceived as helpful for families of CSHCN, a better understanding of the daily life of these children and their families was considered a critical step. Our intent was to identify a small group of parents with extensive experience in advocating for improved care for their children, and who were passionate about improving systems of care for other families.

Five parents were ultimately identified. The chronic clinical conditions found in their children included: Down Syndrome (3 children), Celiac Disease (1 child), Cerebral Palsy (1 child), Cystic Fibrosis (1 child), and developmental delays not yet associated with a diagnosis (1 child). Within these families, the mothers were the primary caregivers. Local groups involved with families of CSHCN, such as Utah State University's Center for Persons with Disabilities[1], the Utah Down Syndrome Foundation[2], and the Utah Family Voices[3], provided great assistance in the family selection process.

Each family received a survey designed to help identify their computer literacy and the level of complexity of the care required by their children. The survey was preceded by a brief description of the project, along with phone and email contacts that could be used to obtain further explanations. Table 1 presents the survey questions.

The five families were also invited to participate in a focus group discussion. The focus group discussion was organized as a mediated phone conference call where each family described a typical day in their lives related to their disabled children. Each family was asked to elaborate on the potential value and desired features of an electronic parent-controlled health record. The session lasted almost 2 hours and was recorded after consent was obtained from each participant. Table 2 presents a summary of the topics discussed.

The information obtained from the families was reviewed and annotated by the authors, leading to a comprehensive set of needs and features. A review of the literature pertinent to PHRs, with a special emphasis on systems proposed for children was also performed.

1 http://www.cpd.usu.edu/
2 http://www.udsf.org/
3 http://www.familyvoices.org/

Table 1 – Survey questions distributed to the families

Questions	Answer
How many medical providers is your child involved with?	[Numeric]
How many other types of professionals is your child involved with?	[Text]
Do you have a computer at home?	Yes; No
Do you have Internet access?	Yes-Dial-up; Yes-DSL; No
How often do you use the Internet (World Wide Web)?	Daily; Weekly; Monthly; Rarely; Other (specify)
How comfortable are you with using a computer?	Very Comfortable; Somewhat Comfortable; Not Comfortable; Other (specify)
Do you use a Personal Digital Assistant (PDA) or hand-held organizer?	Yes; No
What type of medical equipment do you use in your home?	Blood pressure device; Oximeter; Nebulizer; Glucometer; CPAP; Other (specify)

Results

A long list of desiderata was generated from the focus group discussions and subsequent interactions with the families and the local groups involved with CSHCN. Many identified features are recognized as traditional PHR features, including the desire to track provider visits, ask questions prior to the next visit, review treatments, labs and tests ordered and medications prescribed, award password protection and proxy rights to designated others, edit the record but with a provider or care-coordinator validating its accuracy, link to patient-friendly medication information, and "make sure" that providers read what was entered in the record.

The sections below highlight features considered important by families of CSHCN, with special emphasis on items that are not described as standard PHR functions [15-17].

Core features

A *PedMHR* should expand on the important universal elements of a personal child health record (e.g., growth charts, developmental monitoring, immunization records, and advice/information [18]) to develop a comprehensive record that will meet the more complex requirements of CSHCN, their families, and their providers of health and other services.

The key features of this record should include: ease of accessing, recording, and organizing information about care; secure and customizable sharing of information with others in electronic or printed formats; focused links to relevant and reliable information and decision support; links to services and other resources, both locally and nationally; integration of voice recording, transcription, and indexing of visits with providers; integrating financial management tools; and tools for creating and maintaining care plans, with reminders for parents and providers.

Table 2 – Topics selected for the focus group discussion

Topic
Please describe a "day in your life", focusing on activities related to your child, either directly or indirectly.
If you could design the perfect system of storage and access to your child's health records, what would you want in that system?
How would you want to receive new information and learn new activities that may be helpful to your child?
What would you want to help you objectively record and analyze the progress your child is making based on the care plan from the doctor?
What would you want to help you and your doctor detect problems that may affect the treatment and therapies that your child is receiving?
Would you be opposed to having your child's medical record on the Internet at a password-protected site?
Would you want a record of your child's appointments, diagnoses, history, medications, providers, etc, as a reminder for you and to hand to new providers so you don't have to repeat your story?
Should it include a way to send messages to your doctor or other providers, or schedule appointments, refill prescriptions, etc?
How about the ability to write down things to remember, notes, questions to ask, progress your child made, etc., and to link that to your provider(s) or to guidelines to help identify problems?
What other things have you thought of today that you want to share with us?

Information acquisition and storage

Conversations during physician visits are highly valued by patients/parents and generally assumed by physicians to effectively transmit their opinions, explanations, and recommendations. However, much of what is said is missed or forgotten by parents after the visit [19], impairing compliance with treatment and depriving family members and other caregivers of the opportunity to understand the "what, why, and how" of the physician's suggestions. The written record, transmitted electronically or in printed format (and later scanned into the electronic record), may be sufficient to communicate the "what", but is not likely to adequately communicate the "why and how." Some patients have found audio recordings of physician visits to be helpful both to review and remember details and to allow others (e.g., family members) to better understand the communications [20].

A *PedMHR* should provide multiple mechanisms to acquire information, including electronic forms for direct data entry, paper scanning into digital documents, and uploading of digital audio and video. However, information captured using these mechanisms has to be properly structured and encoded to enable computerized repurposing and decision support.

The integration of pre-visit questionnaires, screening instruments (e.g., for development, behavior, depression), medication response assessment tools, and treatment diaries into a *PedMHR* should assure access to them, facilitate their completion (via reminder notices), transmission, and evaluation or validation (prompted by an email notice to the provider), and their integration into the ongoing record.

Perhaps the most useful feature of a *PedMHR* for a child with a chronic, complex condition would be a documented care plan. Such plans allow details of care needed by CSHCN to be recorded and shared with parents, providers, and designated others (e.g., family members, educators). Care plans may also include goals against which progress can be measured, timelines able to trigger reminders of needed labs, appointments, or phone calls, and links to information, instructions, and data entry forms.

Information access and reporting (repurposing)

Parents often lament the need, particularly when their child is being admitted to a hospital, to repeat their "story" over and over to interns, residents, attendings, specialists, etc. A *PedMHR* should dynamically provide preformatted and customizable reports to alleviate much of this frustration and assure the accurate transmission of appropriate and current information. Ideally, such reports should be provided electronically and include a rich set of links and annotations that would offer additional details about specific problems, prior reactions to medications, parent-recorded symptom diaries, and other features. These reports should also be configured with detailed utilization monitoring, enabling parents to ascertain who accessed the information and when.

Similarly, templates for "standard reports" such as letters of necessity, plans of care, school forms, insurance reports, etc., should direct the extraction of information from a *PedMHR* and its subsequent customization and formatting. Condition-specific growth charts (e.g. Down syndrome) should plot growth from measurements entered by providers or caregivers.

Links to pertinent information and knowledge resources

A key need identified by parents of CSHCN is access to information about their child's condition, available services and other resources, educational interventions, growth and development, support groups, financial management, and other aspects of caring for their child and family [21-22]. A number of web sites have been developed over the past several years aimed at providing such information for parents, physicians, and others. The Med-

Home Portal[4] has been serving many of these information needs in Utah since 2001. A *PedMHR* should offer configurable context-aware and data-driven "infobuttons" [23] to sources of specialized information, enabling parents to continuously improve their knowledge and understanding of their child's condition. The utilization of these links should also trigger simulated scenarios followed by focused tests, helping parents to practice what they have learned.

Knowledge sharing and collaboration

A *PedMHR* should enhance the ability of parents and physicians to share knowledge, information, and ideas. Parents of children with complex conditions often become experts in those conditions, particularly in their child's manifestations and responses to treatments. A *PedMHR* should enable parents and providers to share valued resources (e.g., web sites, articles, experiences, observations, best practices) using a variety of online tools (e.g., blogs, message boards, podcasts, wikis, vlogs). A *PedMHR* should make it possible for parents and providers to critique existing resources and suggest new ones, and also to allow parents and providers to develop, individually or collaboratively, a wide variety of brand new educational resources. Given the significant effort required to create such educational resources, a *PedMHR* community should let parents and providers receive monetary rewards for their authoring and editorial efforts, with the perceived usefulness of a resource defining its commercial value.

Communication with providers and care givers

A *PedMHR* should provide for asynchronous messaging and consultations among parents, providers, and caregivers. Asynchronous communications between physicians and patients via email and web-based secure messaging is gradually increasing and some insurers have piloted compensating physicians for such services. Though some resist providing care without face-to-face contact, there are many situations in which electronic consultations may be preferable. "Curbside consultations" by primary care physicians with subspecialists provide potential value in saved time and expense for patients, expediting obtaining an opinion or beginning an intervention, and enhancing the knowledge of the primary care provider. Accomplishing such consultation electronically, linked to authorized access to a *PedMHR*, could potentially be more efficient and useful. With integration of demographic records, these messages could also serve as documentation for billing patients or third parties. Compensation would provide added incentive to physicians and acknowledge the value of the service to all parties.

Financial management

For parents of CSHCN, managing family finances and negotiating the maze of insurance benefits can be a frustrating, nearly full-time job. A *PedMHR* should help families by integrating financial management features, including tools for tracking medical bills and payments, links to insurance companies, suppliers, and providers, and providing reports for use with tax preparation software. Similarly, the clinical activities tracked by a *PedMHR* should be used to document and trigger the appropriate charges and reimbursement. Parents should retain complete control over the content and routing of electronic transactions.

4 http://medhomeportal.org/

Integration with local and regional repositories

As communications among health data repositories become more standardized and patient privacy issues are better integrated into information systems, linking a *PedMHR* to those systems and to service sites should serve to further enhance its utility. In Utah, the "Child Health Advanced Records Management" (CHARM)[5] system is compiling personal health information from several state databases, serving as an electronic broker to improve access. Similar efforts should enable a *PedMHR* to share data with vital records, the state's immunization registry, newborn screening and birth defects registries, and other programs' databases. "Utah Clicks"[6] is a web-based "universal application system" enabling parents to apply for multiple services for children, including: Medicaid, "Baby Your Baby" (prenatal support services), "Head Start and Early Head Start", "Baby Watch" (early intervention), and the CSHCN program (evaluation and treatment for developmental and medical problems). A *PedMHR* should seek integration with such systems and repositories, thereby not only providing and retrieving relevant data in real-time, but also using "intelligent" (context-aware) interfaces to find and access the appropriate services.

Discussion

Despite the significant achievements of the informatics community in defining standardized and interoperable systems, substantial efforts are still required to properly structure and encode the information associated with complex clinical conditions, such as those presented by CSHCN. Similarly, automated methods for consistently extracting precise information from a wide variety of textual and non-textual documents also remains an important challenge.

Without the availability of "computable" data most of the advanced information retrieval and repurposing features identified will not be feasible, dramatically reducing the usefulness of a *PedMHR*. The consistent extraction and repurposing of information represented in clinical documents is one of the core features of a *PedMHR*. However, the structure and content of these documents, as well as their metadata and provenance, must be standardized and predictable, otherwise such features become restricted to documents authored locally by *PedMHR* editing tools.

Particularly in the case of care plans, a number of templates have been published and, in working with our participating parents and providers, it is clear that several new "standard" templates that can be customized to fit various clinical diagnoses and life situations will have to be created. Multiple aspects of the challenges just mentioned regarding the structure and content of the information apply directly to care plans, potentially restricting their usefulness and functionality.

Significant advances in how web-based information resources are structured, indexed, and encoded need to be completed before context-aware and on-demand learning can become a reality. Most of these features are encompassed by what is described as the "Semantic Web"[7]. Our

5 http://charm.health.utah.gov/
6 https://utahclicks.org/index.cfm?
7 http://www.w3.org/2001/sw/

group was recently awarded three years of funding from the National Library of Medicine to substantially expand the MedHome Portal offerings and features, making it compatible with these innovative learning efforts. In terms of web-based communication and collaboration, while most of the technologies mentioned (e.g., blogs, podcasts, wikis) are currently available, user literacy and access must be addressed if all members of the *PedMHR* communities are to become active and knowledgeable collaborators.

A successful implementation of a *PedMHR* clearly has to rely not only on ubiquitous secure communications, but also on advances related to privacy and data security, and technology access. Mechanisms for recognizing and compensating providers for electronic messaging and teleconsultations have to be widely implemented. Potential liability issues must also be addressed. Similarly, ongoing "e-commerce" standardization efforts will greatly benefit the integration of a *PedMHR* with other non-clinical systems. Substantial progress on the specification and standardization of real-time interoperable web services are required before seamless integration can be obtained.

Conclusion

We have attempted to identify the core features of a special kind of PHR, characterized as a *PedMHR*. In essence, a *PedMHR* would empower parents of CSHCN to effectively obtain, organize, understand, and communicate the information necessary to help their children receive the best possible care. This vision was promptly endorsed by the families of CSHCN we had privilege to interact with and also by their providers. Our intent is to eventually create a *PedMHR*, where the parents/family have complete control over the clinical data and can use these data not only to ensure optimal care planning and coordination, but also to acquire and disseminate knowledge about the conditions and needs of their children. We are certain that CSHCN, along with their families and providers would greatly benefit from a *PedMHR*.

Acknowledgments

We gratefully acknowledge Gina Pola-Money and Jefferson C. Sheen, MSW for their help in conceptualizing a *PedMHR* and in organizing the parent focus group.

References

[1] van Dyck, P. C., Kogan, M. D., McPherson, M. G., Weissman, G. R., & Newacheck, P. W. (2004). Prevalence and characteristics of children with special health care needs. *Arch Pediatr Adolesc Med*, 158(9), 884-890.

[2] McPherson, M., Arango, P., Fox, H., Lauver, C., McManus, M., Newacheck, P. W., et al. (1998). A new definition of children with special health care needs. *Pediatrics*, 102(1 Pt 1), 137-140.

[3] Newacheck, P. W., & Halfon, N. (1998). Prevalence and impact of disabling chronic conditions in childhood. *Am J Public Health*, 88(4), 610-617.

[4] Newacheck, P. W., Stein, R. E., Bauman, L., & Hung, Y. Y. (2003). Disparities in the prevalence of disability between black and white children. *Arch Pediatr Adolesc Med*, 157(3), 244-8.

[5] Chernoff, R. G., Ireys, H. T., DeVet, K. A., & Kim, Y. J. (2002). A randomized, controlled trial of a community-based support program for families of children with chronic illness: pediatric outcomes. *Arch Pediatr Adolesc Med*, 156(6), 533-539.

[6] Farmer, J. E., Clark, M. J., Sherman, A., Marien, W. E., & Selva, T. J. (2005). Comprehensive primary care for children with special health care needs in rural areas. *Pediatrics*, 116(3), 649-656.

[7] Liptak, G. S., Burns, C. M., Davidson, P. W., & McAnarney, E. R. (1998). Effects of providing comprehensive ambulatory services to children with chronic conditions. *Arch Pediatr Adolesc Med*, 152(10), 1003-1008.

[8] Palfrey, J. S., Sofis, L. A., Davidson, E. J., Liu, J., Freeman, L., & Ganz, M. L. (2004). The Pediatric Alliance for Coordinated Care: evaluation of a medical home model. *Pediatrics*, 113(5 Suppl), 1507-1516.

[9] American Academy of Pediatrics (2002). The medical home. *Pediatrics*, 111, 184-186.

[10] Ziring, P., Brazdziunas, D., Cooley, W., Kastner, T., Kummer, M., Gonzalez de Pijem, L., et al. (1999). Care coordination: integrating health and related systems of care for children with special health care needs. *Pediatrics*, 104(4 Pt 1), 978-981.

[11] Sia, C., Tonniges, T. F., Osterhus, E., & Taba, S. (2004). History of the medical home concept. *Pediatrics*, 113(5 Suppl), 1473-1478.

[12] Walton, S., Bedford, H., & Dezateux, C. (2006). Use of personal child health records in the UK: findings from the millennium cohort study. *BMJ*, 332(7536), 269-270.

[13] Menachemi, N., Ettel, D. L., Brooks, R. G., & Simpson, L. (2006). Charting the use of electronic health records and other information technologies among child health providers. *BMC Pediatr*, 6, 21.

[14] *The Value of Personal Health Records - A Joint Position Statement for Consumers of Health Care*. (2006). Retrieved Dec 1, 2006, from http://www.amia.org/inside/releases/2006/ahima-amiaphrstatement.pdf

[15] Kim, M., & Johnson, K. (2002). Personal health records: evaluation of functionality and utility. *J Am Med Inform Assoc*, 9(2), 171-180.

[16] Simons, W., Mandl, K., & Kohane, I. (2005). The PING personally controlled electronic medical record system. *J Am Med Inform Assoc*, 12(1), 47-54.

[17] Tang, P., Ash, J., Bates, D., Overhage, J., & Sands, D. (2006). Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. *J Am Med Inform Assoc*, 13(2), 121-126.

[18] Hampshire, A. J., Blair, M. E., Crown, N. S., Avery, A. J., & Williams, E. I. (2004). Variation in how mothers, health visitors and general practitioners use the personal child health record. *Child Care Health Dev*, 30(4), 307-316.

[19] Kessels, R. P. (2003). Patients' memory for medical information. *J R Soc Med*, 96(5), 219-222.

[20] Liddell, C., Rae, G., Brown, T. R., Johnston, D., Coates, V., & Mallett, J. (2004). Giving patients an audiotape of their GP consultation: a randomised controlled trial. *Br J Gen Pract*, 54(506), 667-672.

[21] Farmer, J. E., Marien, W. E., Clark, M. J., Sherman, A., & Selva, T. J. (2004). Primary care supports for children with chronic health conditions: identifying and predicting unmet family needs. *J Pediatr Psychol*, 29(5), 355-367.

[22] Perrin, E. C., Lewkowicz, C., & Young, M. H. (2000). Shared vision: concordance among fathers, mothers, and pediatricians about unmet needs of children with chronic health conditions. *Pediatrics*, 105(1 Pt 3), 277-285.

[23] Fiol, G. D., Rocha, R., & Cimino, J. J. (2006). *Infobutton Standard API Proposal*. Retrieved Dec 1, 2006, from http://www.hl7.org/Library/Committees/dss/infobutton%2Dproposal%2D2006%2D01%2D04%2Ezip

Address for correspondence

Roberto A. Rocha, MD, PhD
RemedyMD, Inc;
9350 South 150 East - Suite 950;
Sandy, UT 84070; USA;
Email: rrocha@remedymd.com

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Empowering Patients to Improve the Quality of Their Care: Design and Implementation of a Shared Health Maintenance Module in a US Integrated Healthcare Delivery Network

Eric G Poon, MD MPH[1,2,4]; Jonathan Wald, MD MPH[1]; Jeffrey L Schnipper, MD MPH[2,4]; Richard Grant, MD MPH[3,4]; Tejal K Gandhi, MD MPH[2,4]; Lynn A Volk, MHS[1]; Amy Bloom, MPH[1,2]; Deborah H Williams, MHA[2]; Kate Gardner[1]; Marianna Epstein[1]; Lisa Nelson[1]; Alex Businger[1,2]; Qi Li, MD MBA[1]; David W Bates, MD MSc[1,2,4]; Blackford Middleton, MD MPH MSc [1,2,4]

[1]Partners Information Systems, Partners HealthCare, Boston, MA, USA [2]Brigham and Womens Hospital, Boston, MA, USA [3]Massachusetts General Hospital, Boston, MA, USA [4]Harvard Medical School, Boston, MA, USA

Abstract

We describe a health maintenance module within a personal health record designed to improve the quality of routine preventive care for patients in a large integrated healthcare delivery network. This module allows patients and their providers to share an online medical record and decision support tools. Our preliminary results indicate that this approach is well-accepted by patients and their providers and has significant potential to facilitate patient-provider communication and improve the quality of routine health maintenance care. Further research will determine the long term impact and sustainability of this approach.

Keywords:

personal health records, quality of care, preventive care, patient-centered care.

Introduction

In spite of spending more per capita on healthcare than any other country in the world, the United States (US) faces significant gaps in the quality of care patients receive(1). Recent research suggests that the US population only receives about half of the recommended evidence-based care, including routine health maintenance (HM) care such as cancer screening tests and adult vaccinations(2). Not only do these gaps in care endanger patients well-being, they represent lost opportunities to deliver cost-effective care to the population at large.

Several reasons likely account for these gaps in quality. Medical care in the US historically has focused on acute care, and not chronic or preventive care. Providers typically are financially rewarded for the quantity of care they deliver, not the quality of care(1). In the ambulatory setting, these factors pressure providers into scheduling more patients per hour rather than spending more time with each patient to address key care issues. In addition, the typical ambulatory practice relies on each individual clinician to detect areas where a patients routine HM care is out-of-date, an unreliable strategy for improving care given the limitations of human vigilance(3).

More recently, healthcare information technology (HIT) with decision support has been touted as promising means to improve quality in the ambulatory setting. However, many systems in the ambulatory setting do not provide adequate decision support, and even when decision support is available, not all providers use it in their clinical workflow(4). In addition, while research in academic settings has shown that decision support aimed at physicians does lead to increased adherence to preventive care guidelines, the improvements have been variable and at best modest(5-7). Given the challenges busy clinicians face in the ambulatory setting, improvement efforts that are directed solely at them may continue to fall short.

By empowering patients to become active participants in their care, an interactive personal health record (PHR) has the significant potential to overcome these quality improvement challenges(8). If patients are given the opportunity to anticipate the discussions that may occur during the clinical encounter about HM issues, they may be more likely to bring those issues to the attention of busy clinicians. Furthermore, if patients can be more informed about the nature of HM tasks, the patients and their clinicians will be able to discuss these issues more efficiently, thus mitigating time pressure as a barrier to good quality care(9).

At Partners HealthCare (Massachusetts, USA), we have developed and deployed a solution to leverage the power of a tethered PHR(8) to address quality and communication gaps in the deliverance of routine health maintenance care. This manuscript describes the rationale and design of a *health maintenance (HM) module* embedded within our personal health record, and will also present data on how this module is being used and patients attitudes toward it.

Methods

Study setting

Partners HealthCare was formed in 1994 through the financial merger between the Massachusetts General Hospital and Brigham and Womens Hospital. Since its inception, this integrated delivery network has grown to include five community hospitals, four rehabilitation and long term care facilities, and a large network of primary care and specialty physicians. Results of clinical tests performed throughout the network, including those that are involved in routine HM care, are stored in a common clinical data repository (CDR). Clinical activities in the ambulatory setting are supported by the enterprise electronic medical record the Longitudinal Medical Record (LMR) which can access data stored in the CDR in addition to other data types such as vaccinations and clinical notes. There are currently over 7,000 clinical users of the LMR.

Personal health record platform patient gateway

As part of an enterprise-wide strategy to facilitate communication between patients and their physicians, Partners HealthCare began in 2001 to develop a secure patient portal to serve as a communication gateway for participating primary care practices. The product, Patient Gateway (PG), allows patients to renew their medications, review their medication lists, request appointment and referrals, communicate with their practice via secure email, and access a licensed health information library(10). As of November 2006, PG supports 20558 patients across 23 primary care and specialty practices, with more than 4200 unique patients using it in any given month. The functionality described in the remainder of this manuscript was developed as additional features to the base PG product.

Design principles

The design of the HM module within PG was guided by several key informatics and patient-centered care principles:

- Data in the medical record ultimately belong to the patient, and the patients privacy must be protected at all times.
- When patients are given online access to their medical record, they should receive guidance on how to interpret the information stored in their record.
- Successful PHR solutions must facilitate workflow and communication for both patients and clinicians.
- Usability is key to patients and providers acceptance of any informatics solutions.

Functional goals for the HM module

We assembled a team of physicians, informaticians, health services researchers to define how to design the HM module to maximize its likelihood of acceptance by clinicians and patients and its potential to improve health care quality. This multi-disciplinary team set forth the following functional goals:

- To inform patients prior to visiting their primary care physicians about health maintenance care that is appropriate given their age, gender, family history and other co-morbidities.
- To remind patients about HM care items for which they may be overdue, giving them an opportunity to learn why they need them and why they are overdue.
- To inform patients about HM care items for which they are up-to-date and congratulate them.
- To allow patients to initiate the process of updating their HM record if they have had a care item performed at an outside facility.
- To encourage patients to state before the visit how they want to take care of HM care items for which they are overdue, passing that information to the clinician who can use it as a starting point for further discussion and planning.
- To facilitate documentation within the LMR by allowing clinicians to review data submitted by patients and to save them (with modifications if necessary) into the patients medical record.

Defining the knowledge base

In deciding on the types of *patient-centric* reminders the PG HM module would provide, we first looked to prior work that had reviewed well-accepted quality indicators in the US such as HEDIS measures, and trusted sources such as the US Preventive Services Task Force recommendations and other locally accepted clinical guidelines(6). Using the results of that review, a previous project had implemented a set of *clinician-centric* reminders within our electronic medical record LMR to improve clinicians adherence to established HM care guidelines(6). For the current project, we critically reviewed the knowledge base (KB) supporting the LMR *clinician-centric* reminders and, where necessary, updated it. The updated KB addresses the following 3 areas of routine HM care:

- *Womens health*: breast cancer screening, cervical cancer screening, osteoporosis screening
- *Adult Vaccinations*: influenza vaccine, pneumococcal vaccine, tetanus/diphtheria booster vaccine
- *Lipid Assessment*: for patients with average cardiac risk factors, those with diabetes and those with documented coronary artery disease

Colon cancer screening was not addressed in the KB because our clinical data repository did not contain results of endoscopy reports. The absence of these data might cause *clinician-centric* reminders within LMR and *patient-centric* reminders in the HM module to fire inappropriately. We therefore made the decision not to implement colon cancer screening logic at this point.

Because of a separate PG module that also allowed patients and physicians to collaborate on entering coded family history into the electronic medical record, we were able to adjust the recommendations to patients based on their family history. This benefited the logic for breast cancer screening, osteoporosis screening, and lipid management in that high-risk patients would be prompted

in the PG HM module to start screening at an earlier age and at closer time intervals.

Decision support architecture

Beyond allowing patients and clinicians to share information in the medical record, our project allows both parties to share decision support tools to improve the quality of routine HM care. To ensure consistency for the logic behind *clinician-centric* reminders in the LMR and the *patient-centric* reminders in the PG HM module, our team elected to manage the knowledge base (KB) for both types of decision support centrally. As demonstrated in Figure 1, a *reminder engine*, implemented as a series of services that can be used by a variety of applications, determines whether a patient qualifies for a particular HM care item using the *risk group* definitions within the KB. If a patient does not qualify for the HM item (e.g. a male patient in the case of mammography), neither the patient nor the clinician will see a reminder. If the patient does fall into the risk group (e.g. female patient over the age of 50 in the case of mammograms), the reminder engine further determines whether the *overdue condition is* met. If a patient falls under a *risk group* and the *overdue condition* is met (e.g. female patient over the age of 50 who has not had a mammogram within the past year), the PG HM module will inform the patient that the item is overdue and the LMR will inform the provider that the patient needs a mammogram. If, however, the patient falls under the *risk group* but the *overdue condition* is not met (e.g. a female patient over the age of 50 who had a mammogram 6 months ago), the PG HM module will inform the patient that the HM item is up-to-date, and the LMR will not remind the clinician to order a mammogram.

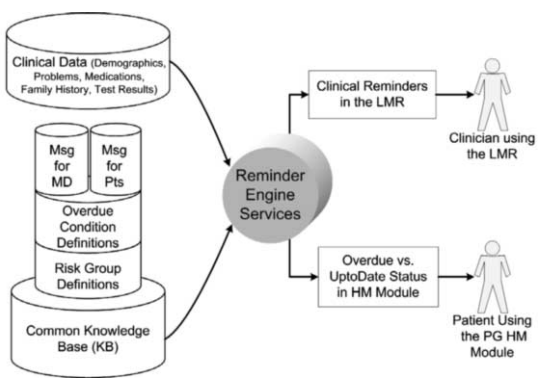


Figure 1 - Clinical Decision Support Architecture

While the patient and clinician may share the same decision support logic, the messages displayed to the two groups are necessarily different: the physician may get a very terse prompt about overdue items while the patient receives more extensive information to help him or her interpret the decision support prompts provided in the HM module. The content for these two groups is maintained separately within the KB.

User interface design process

As the multi-disciplinary team gathered the functional requirements, we iteratively developed a series of low and then high fidelity prototypes to explore different design approaches and refine the details of the design. We also solicited input from an advisory council comprised of primary care physicians from practices that were already using the base PG product. Finally, we conducted a series of usability tests with volunteer patients going through mock scenarios to enhance the ease of navigation and clarity of the user interface.

Functionality review patients experience

Patients with access to PG were invited to participate in this study to evaluate the HM module. Once they completed the consenting process, they were given access to their HM records within the HM module. Figure 2 illustrates the current functionality of this module. HM items that are overdue are grouped together and highlighted. Patients may review, for each overdue HM item, i) a short description of the HM item (e.g. what is a mammogram and why might women need one?); ii) a link to a more detailed description of the HM item, and iii) an explanation of why the patient may be overdue for this HM item (e.g. because the patients age and risk factor call for an annual mammogram and her last one was more than 2 years ago).

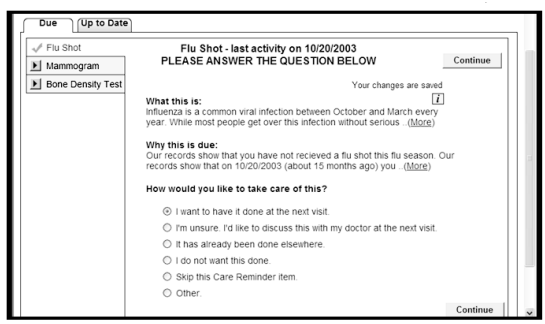


Figure 2 - Patients Experience in PG HM Module

Three weeks before a scheduled visit, patients are sent a secure message in which they are invited to review their HM record online. At this point, patients are prompted to indicate how they may want to address each overdue item in their HM record during the upcoming visit. Patients may indicate one of the following: i) they want to take care of this item at the next visit; ii) they are unsure what to do and would like to discuss it at the next visit; iii) they state that they have had this item taken care of elsewhere, after which they will be prompted to enter the specifics so that their clinician can update the medical record; iv) they do not want to address this item, or v) they want to defer the decision. After reviewing one or more items in their HM record, the patient can view a summary of their desired actions and submit the data for review by their clinician at the upcoming visit.

Functionality review clinicians experience

After a patient has submitted information through the HM module before the visit, the physician is prompted visually during the visit that data have been submitted by the patient. Upon reviewing the health maintenance screen within the electronic medical record (a commonly-used screen regardless of whether patient has used the HM module), data submitted by the patient through the PG HM module, if present, become visible (Figure 3). This information serves as a starting point for further discussion regarding each care item that may be out-of-date, and the information entered by the patient can facilitate documentation. For example, if the patient indicates that a vaccine was given outside the system and enters the date the vaccine was given, that information can be easily saved into the patients medical record once the clinician has reviewed it.

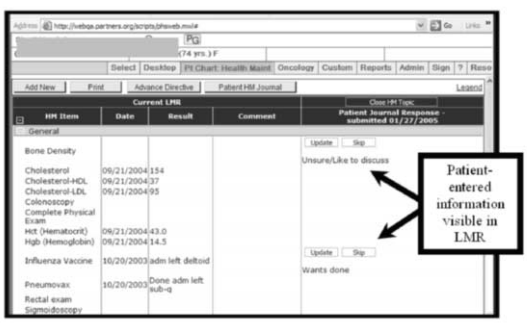


Figure 3 - Clinicians Experience in the LMR

Deployment strategy

We relied on several approaches to deploy the HM module within the PG product platform. First, we identified physician champions in each practice and solicited their input for the optimal ways to deploy the HM module within their practice. Second, we presented the HM module to physicians at practice meetings to introduce its functionality to the clinicians. Third, we provided on-site and online support, and returned to practices to hear feedback about the HM module after rollout. Fourth, we provided marketing materials to practices so that patients became aware of this new feature. Fifth, we sent reminders through the base PG product to all patients who had consented to the study to prompt them to review and update information in the HM module prior to the visit.

Short-term evaluation of the HM module

The HM Module was rolled out over the course of 14 months between October, 2005 and November, 2006 to 7 primary care practices within our integrated delivery network in conjunction with a family history module[1]. To evaluate the usage of the HM by patients and clinicians, we recorded the following: i) the number of patients who accessed the HM module to review their HM records; ii) the number of patients who were invited to update their record prior to their scheduled visit; iii) the number of patients who submitted their information to the clinicians, and iv) the number of times clinicians electronically reviewed data submitted by the patients through the HM module.

To evaluate patients attitudes towards the use of the HM and family history modules, patients who were invited to review their data prior to a scheduled visit were invited to complete a short online survey 3 days after the scheduled visit. On this survey, patients were asked to rate on Likert-scales their experience in reviewing and updating their medical records online, including whether they found the HM and family history modules easy to use, whether they felt more prepared for the visit, and whether the use of the modules led their providers to have more accurate clinical information. Patients were also invited to express other comments in free-text.

Long-term evaluation plan for the HM module

To evaluate the impact of the HM module on quality of care, we are conducting a cluster randomized controlled trial (RCT) in which the 7 practices whose patients are granted access to the HM and family history modules (intervention group) will be compared with 7 practices whose patients are granted access to other modules related to medication management and diabetes care (control group). Clinical outcomes to be assessed in the RCT include adherence to established HM guidelines. Patients knowledge about HM guidelines will also be assessed with surveys administered before and after the intervention period in both the intervention and control arms. We will also assess physicians and their staffs attitudes towards the HM module through separate clinician surveys. Results of these evaluation efforts should be available by late 2007.

Results

Between July 2005 and November 2006 patients who sought primary care at one of the 7 study practices were invited to participate in the study. Of them, 2,779 completed the consenting process, which included a baseline attitude survey. Of the 2,779 patients who consented, 63% were female. Their mean age was 47.4 at the time of consent.

Of these 2,779 patients who completed the consent process, 2,361 (85%) reviewed their HM records. In addition, 970 of 2,779 (35%) patients had a routine visit scheduled at least three weeks in advance and were therefore invited to update their HM record and state their preference for taking care of overdue HM care items. Of these 970 patients, 696 (72%) completed the review and updating process and submitted the information for their clinicians to review. Clinicians reviewed the data electronically within the electronic medical record for 460 (66%) of these patients.

Between July and November 2006, 437 patients who opened their invitation to update their HM record prior to a visit were further invited to respond to an online survey to assess their experience with the HM module. Overall, 179

1 The details of the family history module are being described in a separate manuscript.

patients (response rate = 41%) responded, and 81% of the respondents found the journal very easy or easy to complete. 51% of respondents either strongly agreed or agreed that the use of the journal led their providers to have more accurate information, with another 37% feeling neutral about whether the use of the journal had an impact in this area; 48% of respondents agreed that they felt more prepared for the visit, with another 41% feeling neutral.

While patients response to the HM module was largely positive, review of the qualitative comments from the survey revealed that not all clinicians were aware of their ability to review HM data submitted by the patient. Some patients felt that their clinician did not review data submitted by them. Others reported that their physicians asked them to fill out another paper survey in the waiting room that largely duplicated their interaction with the HM module. Certain patients desired the ability to enter information in free-form to the physician, such as reason for the visit or other active concerns.

Discussion

Our preliminary findings demonstrate that sharing the medical record and decision support tools between patients and their providers is a promising approach for improving quality of care. While not all eligible patients and clinicians used this new set of tools when offered the opportunity, those who did generally found the toolset easy to use and many thought that it made them feel more prepared for the visit and allowed their clinicians to have more up-to-date information.

There are several potential limitations to this approach to improve the quality of routine HM care. First, this approach does not benefit patients who do not have internet access or who do not have sufficient computer literacy to use our system. However, this concern is partially mitigated by the fact that internet access in the US is rapidly increasing (11). In addition, we have taken every effort through usability testing to ensure that our system is easy to use. Second, in our current implementation, only patients who are scheduled to have a visit are invited to update their HM records, and those without upcoming visits do not benefit from this aspect of the HM module. Also, patients are only invited to update their records in focused and structured ways, and cannot express their thoughts and concerns in free-form. We would have liked to provide these features, but physician practices were concerned early on during the project that allowing patients to submit updates to their HM record without an upcoming scheduled visit or providing for free-form entry might impose undue workflow and liability burden on the practices. Further research is therefore needed to examine these issues that arise from the deployment of tethered PHRs. In addition, if our long-term evaluation demonstrates a clinical benefit to the use of the HM module, then it may be possible to articulate the business case for the broader deployment of this approach. Finally, successful deployment of this approach requires the full support of local clinical leaders and extensive training. While we invested significant resources in these areas, our efforts did not reach all clinicians, as evidenced by the fact that some clinicians were not aware that they could review HM data submitted by patients. Optimal ways for deploying this technology at various types of institutions deserve further investigation.

Conclusions

We have implemented a novel approach to improve the quality of routine HM care by allowing patients and their providers to share the medical record and decision support tools. Our preliminary results indicate that this approach is accepted by patients and their providers. Further research will determine the long-term impact and sustainability of this approach.

Acknowledgements

The authors would like to thank the US Agency for Healthcare Research and Quality (Rockville, MD) for supporting this work via grant R01 HS 13326-03.

References

[1] Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.

[2] MCGlynn EA, Asch SM, Adams J, et al. The Quality of Health Care Delivered to Adults in the United States. NEJM 348(2635), 2645. 2003.

[3] Reason J. Organizational accidents: The management of human and organizational factors in hazardous technologies. Cambridge, England: Cambridge University Press; 1997.

[4] Simon SR, Kaushal R, Cleary PD, et al. Correlates of Electronic Health Record Adoption in Office Practices: A Statewide Survey. JAMIA. 2006;in press.

[5] Tierney WM, Overhage JM, Murray MD, et al. Effects of Computerized Guidelines for Managing Heart Disease in Primary Care: A Randomized, Controlled Trial. JGIM. 2003;18:967-76.

[6] Sequist TD, Gandhi TK, Karson AS, et al. A randomized trial of electronic clinical reminders to improve quality of care for diabetes and coronary artery disease. JAMIA 12(4), 431-437. 2005.

[7] Shea S, DuMouchel W, Bahamonde L. A meta-analysis of 16 randomized controlled trials to evaluate computer-based clinical reminder systems for preventive care in the ambulatory setting. JAMIA 1996;3:399-409.

[8] Tang PC, Ash JS, Bates DW, et al. Personal Health Records: Definitions, Benefits, and Strategies for Overcoming Barriers to Adoption. JAMIA 13(2), 121-126. 2006.

[9] Markle Foundation and Robert Wood Johnson Foundation. Connecting Americans to Their Healthcare. 2004.

[10] Poon EG, Wald J, Bates DW, et al. Supporting patient care beyond the clinical encounter: three informatics innovations from Partners Healthcare. Proceedings, Ann AMIA Symp, 1072. 2003.

[11] U.S.Census Bureau. Computer and Internet Use in the United States: 2003. U.S.Census Bureau . 2005.

Address for correspondence:

Eric Poon, MD MPH;
Email: epoon@partners.org

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Clinical Communication Ontology for Medical Errors

Yang Gong[a], Min Zhu[b], Jun Li[b], James P. Turley[b], Jiajie Zhang[b]

[a]*School of Medicine, University of Missouri-Columbia, 65211, USA*
[b]*School of Health Information Science, University of Texas Health Science Center at Houston, 77030, USA*

Abstract

Clinical communication failures caused 60% of sentinel events reported by the Joint Commission on Accreditation of Healthcare Organizations. The difficulties of communication have been the primary cause of errors leading to patients' death. For analyzing medical error events, uncovering the patterns of clinical communication, this paper reports the design and development of clinical communication ontology. The ontology contains eight axes and was validated using ten medical error cases, where communication was the main factor. The coding process demonstrates that the ontology can be used as a guideline for future medical error reporting system, through which the root cause of medical error due to communication will be revealed in a clear pattern. This ontology contributes to the generation of proper interventions and effective strategies for reducing medical errors.

Keyword:

medical error, communication, ontology, cognitive factors

Introduction

According to the Institute of Medicine, medical errors killed 44,000 to 98,000 people each year in the United States, which is more than breast cancer or highway accidents[1, 2]. Improper clinical communication is nearly twice as common as improper skills in causing preventive injury or death[3].Widespread attention to medical error and its prevention is becoming part of the culture in most health care organizations.

Researchers have found that the causes of errors are complex, involving human slips and mistakes and that systems weaknesses that predisposed to human error are important and recommended checks[4]. It is not surprising that big mistakes can lead to bad outcomes. However, unnoticed and un-rectified small mistakes, when they are accumulated, can cause bad outcomes as well. In this situation how errors occur is similar to a game, called *whispering game* or *telephone*, where a line of people are given a message at one end which is whispered to the next person down the line until it gets to the other. The final recipient then announces the phrase that he or she has heard, and it is compared to the original. Most of the time, the original message is distorted due to the cognitive distance, such as the difference of memory, language skill, and attention, etc. among the people in the communication line.

The *whispering game/telephone* phenomenon is a basic face-to-face communication, falls into the core level of the system hierarchy illustrating cognitive factors in medical errors [4], where the individuals trigger errors, and no technology usage is involved. Beyond this core level, individuals and technologies in clinical communication such as telephone, patient chart, whiteboard, post-it sticker, etc interact with each other; therefore, can cause more complex errors. At the level of distributed systems, communication is more complex among groups of people and technologies used by the groups. For example, it is prone to errors due to interaction when moving a patient to another unit or to a new team during a shift change. A flubbed handoff in 1995 caused one man in Florida having the wrong leg amputated[5].

Clinical communication is the main contributing factor to medical errors. Problems related to communication have taken 60% of sentinel events reported to the Joint Commission on Accreditation of Healthcare Organizations[6]. The information exchanged through such communication negatively affects the quality of decision making as well as healthcare quality. Earlier in 2006, in response to this report, the Joint Commission on Accreditation of Healthcare Organizations released a draft requiring U.S. hospitals to standardize their approach for handoff communications or risk losing their accreditations.

Clinical communication is carried out in various patterns, which can be involved with or without technologies. Interactions during communication can occur between person and person; person and technology; technology and technology. Observational studies have revealed that clinicians prefer face-to-face communication which is supposedly efficient in most occasions but frequently causes interruption and high communication loads[7-10]. The high workload with an interruptive nature in multitask setting may negatively affect clinicians' memory, attention, and work performance and causes medical errors. The difficulties of communication have been the primary cause of errors leading to the death[11]. Researchers have identified that clinicians spend 80% of their times in communication, with 30% of all communication events classified as interruptions[9]. Therefore, improving clinical communication efficiency and quality is helpful and significant in reducing medical errors.

Although researchers [12, 13] in different domains have classified various types of communication errors that occur between humans, and between humans and comput-

ers, little effort has been made to identify the cognitive factors which cause miscommunication clinical culture rather than device malfunction, channel failure, and low quality of signal, etc. Medical errors are products of cognitive activities when situated in a multitask, time critical and high workload domain[14]. Since medical errors are often triggered by human errors that occur during teamwork[4], the quality of communication among the team members has become our research interest.

In the context of communication, individuals, technologies, cognitive factors, and other factors can be used as key concepts to describe how communication is carried out. The communication ontology provides formal definitions and coverage of various concepts relevant to the domain of communication. It can be used to develop a knowledge-based reporting system through unification of information from data across various organizations, practice domains, and applications, and have sufficient detail to be of practical use. This ontology is an important component of any medical error ontology.

The purpose of this article is to describe the development of ontology of clinical communication for medical errors and explain how we use the ontology to support collection, storage and interpretation of the communication activities in clinical settings.

Materials and methods

Communication error ontology

An ontology can be defined as a specification of a conceptualization[15], it defines a common and controlled vocabulary for the purpose of enabling knowledge sharing and reuse of information[16]. Ontology has a broader application and richer functionality than a taxonomy which is a collection of controlled vocabulary terms organized into a hierarchical structure. A formal ontology is a controlled vocabulary expressed in an ontology representation language, for example OWL. This language has a grammar for using vocabulary terms to express something meaningful within a specified domain of interest. The grammar contains formal constraints on how terms in the ontology's controlled vocabulary can be used together.

To the best of our knowledge, there is no ontology on medical errors due to clinical communication. We have successfully retrieved some observational studies in the medical field on the medical errors due to communication [4-6] and a prototype of communication error taxonomy in the technology field [9]. These are of great help in collecting key concepts used in the ontology.

Taking grounding as a theoretical framework, we propose the communication ontology covering three levels of cognitive factors which include Level I, individuals; Level II, Individual-Technology Interaction within a team; Level III, Individual-Technology Interaction among teams. The collection of concepts included in the ontology was based on observational results of clinical activities and existing communication error taxonomy.

Although there are no methodologies of building ontology mature enough and widely accepted. The main methodology we employed to build ontology is ENTERPRISE methodology[17], which proposes the following stages:

1. identify the purpose and scope of the ontology;
2. build the ontology by capturing knowledge, coding knowledge and reusing appropriate knowledge from existing ontology;
3. evaluate the ontology;
4. document the ontology.

This methodology contains a set of techniques, methods and guidelines for each stage. Examples are, during capturing knowledge stage, identifying key concepts and relations, producing unambiguous text definitions for such concepts and relationships, identifying terms to refer to such concepts and relationships via brainstorming and meetings with domain experts.

Ontology scope and requirements

The concepts we collected for developing the ontology have been designed to satisfy Cimino's desiderata for controlled medical vocabularies[18], which are the core concepts in our ontology. Cimino has articulated that medical controlled vocabularies should include the following: Comprehensive content, concept-based, formal definitions, concept permanence, Multiple hierarchies, meaningless concept identifiers, do not use "not elsewhere classified", multiple granularities, multiple consistent views, context specific information, graceful evolution, and composition-decomposition.

An ideal ontology can introduce a host of structural and conceptual relationships including superclass/ subclass/ instance relationships, property values, time relationships, and others depending on the representation language used. As a collection of communication concepts in clinical settings, first step was to build up a thesaurus, and define each concept and the relationships. The collection of concepts is expected to be exhaustive and mutually exclusive. To be exhaustive is to reach the full coverage of the concepts needed for the description in clinical communication. To be mutually exclusive is to maximally reduce the conceptual overlapping between concepts used in the ontology.

Development approach

Knowledge acquisition was conducted through brainstorming and meetings with domain expert on the basis of literature review which provides multiple observational results on clinical communication[8, 9, 19] and some communication taxonomies for other domains.

Conceptual model of the ontology was developed by employing a middle-out approach. Instead of a bottom-up or top-down approaches, the middle-out approach begins by conceptualizing and defining the concepts that are more highly connected to other concepts. The definitions of core concepts, therefore, can be used as references to build up the definitions of simpler concepts. By doing so, it allows the use of an ontology reasoner to check whether or not all of the statements and definitions in the ontology are mutually exclusive and consistent, and to maintain a correct hierarchical structure of the conceptual model.

Protégé-OWL editor was used to build the ontology. RacerPro was used as a reasoner to check the definitions and

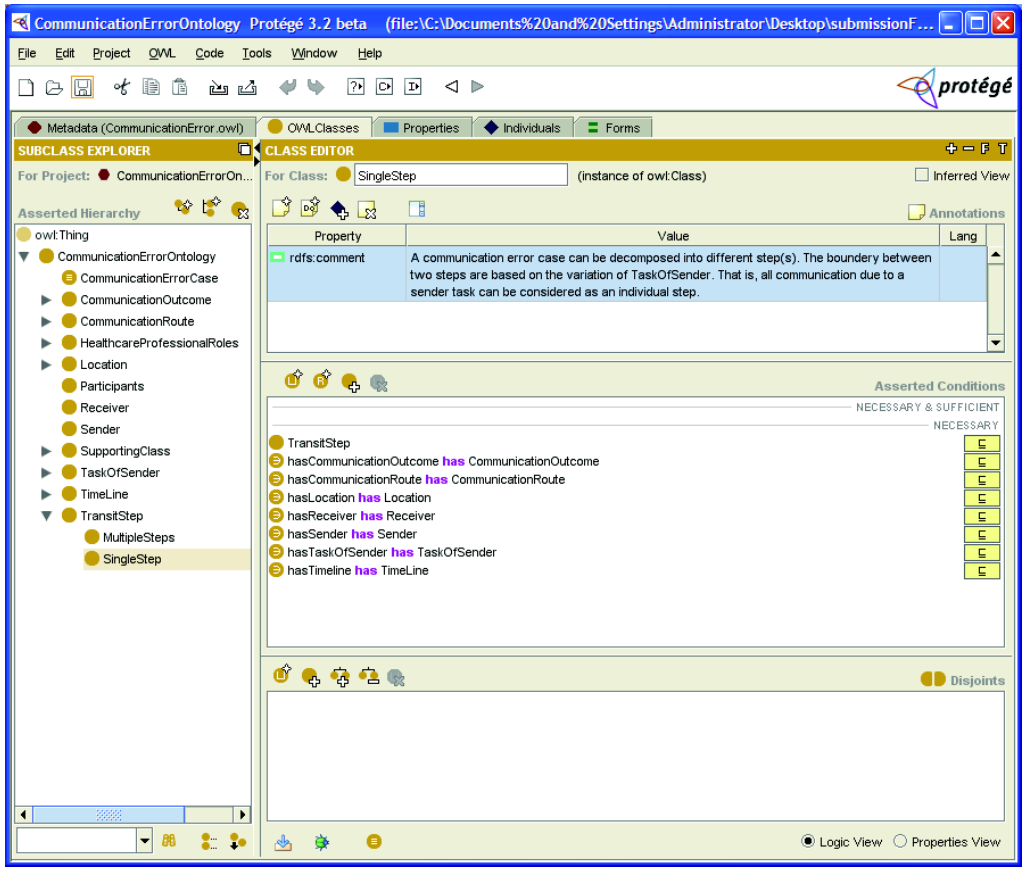


Figure 1 - CommunicationErrorCase Class presented in Protégé 3.2 Beta

logics in the ontology. The core part is the CommunicationErrorCase class which models a medication error due to communication. All the other classes are either the properties of the CommunicationErrorCase class or supporting class for these properties. In the ontology, a communication case typically contains steps by which information is communicated among senders and receivers and patterns associated with the steps. Two levels of classes were used in the ontology to describe the information relay in specific steps and the overall case pattern in general.

At the step level, each step is attached a time when the step happens, a location where the error occurs, the sender(s) or receiver(s) who involved in the activity. It also includes the description of messages which is communicated, the media used to carry the content of communication, and an outcome for each communication step. The CommunicationOutcome class is further defined with four levels of coordination for grounding mutual understanding. The levels are Conversation, Intention, Signal, and Channel. Each level is built on top of its lower level. The communication error could occur in any of these levels. In the situation where a communication error occurs at a level, none of its higher levels can be complete.

At the case pattern level, overall communication patterns are built in such as single step, linear step, circuit step, hybrid of linear and circuit step and so on. Communication participants are identified according to their roles. For example, clinicians usually play roles of sender in one step and may become receivers in the other. Patient outcomes are defined and incorporated with the standard medical nomenclature or medical error ontology[20].

As a result of such consideration, totally eight axes have been created for the ontology, which are TimeLine, Location, Participants, MessageCharacteristics, Media, TaskOfSender, TransitStep, and CommunicationOutcome classes. *Figure 1* shows the eight classes presented in Protégé 3.2 beta version. In each class related to the communication error case, we defined each class by adding subclasses, properties, and restrictions. The example of how to code communication error cases by our ontology is demonstrated in *Figure 2*.

Case coding as a process of validation

As a result of online search and literature, we collected 250 medical error cases for examining the usability of the communication ontology. Authors scrutinized all the cases and identified the cases reported communication as the cause of medical errors.

In general, the authors coded cases using the ontology in the following steps:

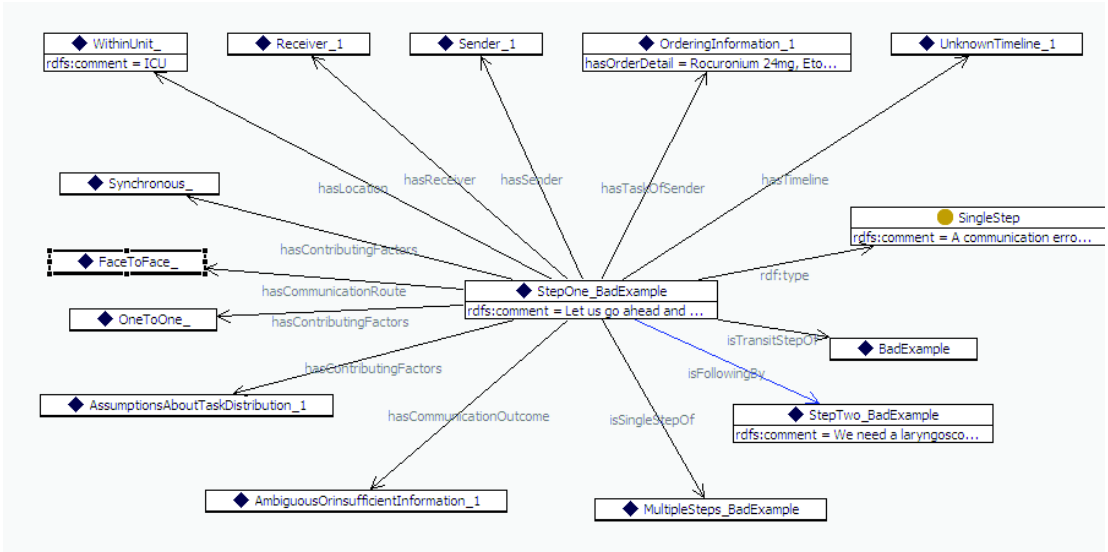


Figure 2 - Case Coding Results

1. Selecting communication from the medical error cases we had collected.
2. Selected communication cases were analyzed and coded individually by authors based on previous identified classes in protégé.
3. Three authors then discussed the case content, granularity and possible assumptions.
4. One of the authors then coded the case into protégé as final version. Meantime, we upgraded the ontology towards a more exclusive and exhaustive version as well.

Figure 2 shows the results of coding a case occurred in ICU where a group of clinicians worked collaboratively on a patient of respiratory failure. An attending physician tried to place an order with Rocuronium 24mg, Etomidate 10mg and atropine 0.4mg. Unfortunately, he did not specify the time point when these medications should be admitted. Shortly after his verbal order, an inexperienced nurse admitted the medication directly before a respiratory therapist had prepared the patient's respiratory track. As a result, the patient lost his spontaneous breath due to the medication admitted in an incorrect time point. In response to this emergency, the attending physician and respiratory therapist placed an artificial airway immediately. The entire case was decomposed into 16 steps based on the communication ontology classes. The 15 steps are coded as success and the very first step was defined as a communication error due to unspecified medication admission.

Results

Using the ontology, we are able to code all of the selected medical error cases due to communication. However, the information described in the reports has various levels of granularity. We identified two patterns of case reports while analyzing 10 cases. In one pattern, cases were coded successfully by making minimum assumptions. Under this pattern there were two subcategories. For the case with a detailed description, there was no need to get hospital-specific information. For some cases, coders had to be familiar with the hospital culture and procedures without which coders could not complete the coding tasks. In the other pattern, cases were coded with an amount of assumptions based on coder's understanding of healthcare domain. There were some cases that could not be coded due to the lack of detailed information. This detailed information is expected during the reporting of medical errors. The reporter of the medical error event could be aware of this missing information, if the ontology had guided the reporter to report the incident in a complete and useful manner. An ontology driven reporting system has the advantage in collecting the expressive and mandatory information. In return, the collection of information guided by ontology will play an important role in analyzing error cases, revealing communication patterns and reducing reoccurrence of the same kinds. The cases cover a variety of methods of communication used in the clinical workflow. Analyses of the cases indicate that clinicians favor the synchronous form of communication even though this kind of communication is highly interruptive to their workflow [8]. This result is consistent with other researcher's observational studies.

Discussion and future work

This paper describes our experience of developing clinical communication ontology to serve as a means of describing clinical communication error, a subset of medical errors. Concepts, relationships and properties and so on in the ontology are subject to modification during analyzing more communication cases.

The ontology primarily provides eight axes for description of the clinical communication. It allows the users to categorize the communication activities and ultimately figure out the patterns for effective communication among clinicians. The development is an iterative process, the

modification and reconstruction will be continuous during the case coding process. We aim at an exhaustive and exclusive ontology in clinical communication which not only describes the activities but also be able to analyze communication pattern, explain communication error, and eventually suggest proper solutions.

We noticed that there is a tradeoff between the ontology and the case description. We tried to balance these two components to ensure that the ontology was exclusive and exhaustive. Based on the case description we did not make any major changes to the ontology and did not make too many assumptions to make the case adequate for coding. This implies that the ontology can be tailored into different granularity fitting the case reporting requirement.

One limitation of this study was that the medical error cases collected were only a small sample of all events and hence they do not sufficiently represent the population. We would require a larger sample from a healthcare quality reporting system in practical use to ensure that the cases adequately represent the population.

Conclusion

The communication ontology for medical error has its practical implication in improving patient safety and reducing medical error. One way to learn from errors is to establish a medical error reporting system, where medical error data are collected in a structured format and can be useful for the detection of patterns, discovery of underlying factors, and generation of solutions. Such ontology serves as a guideline for hospital reporting medical errors. It can be customized into different scales so as to meet hospital needs and minimize the burden of data entry in health care organizations.

This ontology will be served as a foundation for analyzing the communication activities within medical errors, revealing the communication patterns and providing information to possible intervention strategies. The ultimate goal is to improve the communication quality and efficiency and thus reduce medical errors. Validating the ontology through medical error cases gives us an insight into the usability and validity of such ontology. Possible modifications to the ontology can also be made based on the findings of such a validation. It would also help us to identify critical missing components of medical errors that are expected to be reported towards a more useful reporting system.

Acknowledgments

This project was supported by Grant R01 LM07894 from the National Library of Medicine. The authors express heartfelt thanks to Julianna W. Brixey, Anwar M. Sirajuddin for their contribution in the design and development of the ontology.

References

[1] O'Connor E. Medical errors kill tens of thousands annually, panel says. CNNcom. 1999 1999-11-29.
[2] Kohn LT, Corrigan JM, Donaldson MS. To err is human. Washington, DC: National Academy Press 1999.
[3] Zinn C. 14,000 preventable deaths in Australian hospitals. BMJ. 1995;310(1487).
[4] Zhang J, Patel VL, Johnson TR, Shortliffe EH. A cognitive taxonomy of medical errors. Journal of Biomedical Informatics. 2004;37(3):193-204.
[5] Naik G. A hospital races to learn lessons of ferrari pit stop. The wall street journal. 2006 2006-11-14.
[6] Unknown. Sentinel Event Alert, June 23, 2003. *Sentinel event statistics*: Joint Commission on Accreditation of Healthcare Organizations 2003.
[7] Bailey LL, Willems BF, Peterson LM. The effects of workload and decision support automation on enroute r-side and d-side communication exchanges. Oklahoma City, OK 73125: Civil Aerospace Medical Institute Federal Avaition Administration; 2001 Dec 2001.
[8] Coiera E, Tombs V. Communication behaviours in a hospital setting: an observational study. bmj. 2005;1998(316):673-6.
[9] Coiera EW, Jayasuriya RA, Hardy J, Bannan A, Thorpe MEC. Communication loads on clinical staff in the emergency department. MJA. 2002;176:415-8.
[10] Parker J, Coiera E. Improving clinical communication: a view from psychology. Journal of the American Medical Informatics Association. 2000;7(5):453-61.
[11] Lingard L, Espin S, Whyte S, Regehr G, Baker G, Reznick R, et al. Communication failures in the operating room: an observational classification of recurrent types and effects. Qual Saf Health Care. 2005(12):330-4.
[12] Paek T. Toward a taxonomy of communication errrors. ISCA Workshop on Error Handling in Spoken Dialogue Systems; 2003; 2003. p. 53-8.
[13] Te'eni D, Sagie A, Schwartz DG, Zaidman N, Amichai-Hamburger Y. The process of organizational communication: a model and field study. IEEE transactions on professional communication. 2001;44(1):6-20.
[14] Zhang J, Patel, VL., Johnson, TR., Turley, JP. . Health informatics and medical error. . Business Briefing: US Healthcare Strategies. 2005:34-5.
[15] Gruber T. A translational approach to portable ontologies. Knowledge Acquisition. 1993;5(2):199-220.
[16] Turley JP, Brixey JJ, Johnson TR, Mokkarala P, Zhang J. Comprehensive medical error ontology for the codification of published literature. Bulletin of the Japanese Cognitive Science society. 2006 March;13(1):1-11.
[17] Uschold M. Building Ontologies: Towards A Unified Methodology. *Proc Expert Systems 96*. Cambridge 1996.
[18] Cimino JJ. Desiderata for ControlledMedical vocabularies in the Twenty-FirstCentury. *IMIA WG6 Conference*. Jacksonville,Florida 1997.
[19] Coiera E. When conversation is better than computation. Journal of the American Medical Informatics Association. 2000;7(3):277-86.
[20] Mokkarala P, Brixey, JJ., Johnson, TR., Patel, VL., Zhang, J., & Turley, JP. Development of comprehensive medical error ontology. Manuscript submitted for publication. in review.

Address for correspondence

Mailing Address: 404 Clark Hall,
Department of Health Management and Informatics,
School of Medicine, University of Missouri-Columbia,
Columbia, Missouri, 65211, USA
Email: gongyang@missouri.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

u-SHARE: Web-based Decision Support / Risk Communication Tool for Healthcare Consumers with Unruptured Intracranial Aneurysms

Aoki N [a, b, c], Sakai M [d], Nakayama T [e], Fukuhara S [d], Ohta S [a, b, f], Kikuchi N [a, b, g], Oishi M [a, h], Kiuchi T [c], Nozaki K [i], Hashimoto N [i]

[a] *Center for Health Service, Outcomes Research and Development – Japan (CHORD-J), Tokyo, Japan*
[b] *School of Health Information Sciences, University of Texas Health Science Center at Houston, Houston, TX, USA*
[c] *University Hospital Medical Information Network (UMIN) Center, The University of Tokyo Hospital, Tokyo, Japan*
[d] *Department of Epidemiology and Healthcare Research, Kyoto University School of Public Health, Kyoto, Japan*
[e] *Department of Health Informatics, Kyoto University School of Public Health, Kyoto, Japan*
[f] *Okayama Central Hospital, Okayama, Japan*
[g] *Palliative Care Service, Tohoku University Hospital, Sendai, Japan*
[h] *Oishi Clinic, Kyoto, Japan*[i] *Department of Neurosurgery, Kyoto University Faculty of Medicine, Kyoto, Japan*

Abstract

Purpose: Clinical management for unruptured intracranial aneurysms (UIA) is controversial and requires professional knowledge which is the main reason that patients have difficulty in making decisions. The purpose of this study is to develop a tool that aids healthcare consumers in making optimal shared decisions with decision analysis.

Methods: The decision model and relevant data were derived from published literature. A web-based decision analytic tool was designed to provide a systematic guide for patients to understand favorable treatment options, intrinsic uncertainty, and critical factors for decision making. Twenty-nine testers evaluated content appropriateness, usability and clinical usefulness of the tool.

Results: The decision analytic tool has been successfully implemented and evaluated. Testers generally judged the web-based decision analytic tool as functional and useful. Acceptance rate for decision analysis was higher in non-healthcare professionals than health care professionals.

Conclusions: Our decision analytic tool was well accepted especially by healthcare consumers. The tool enables UIA patients to enhance their knowledge and understanding toward optimal shared decision making and can be an alternative "structured informed consent tool".

Keywords:

decision analysis, unruptured intracranial aneurysms, decision support tool, health communication, risk communication

Introduction

In spite of the recent prospective multi-center cohort study conducted by the International Study of Unruptured Intracranial Aneurysms Investigators (ISUIA) [1], optimal management for unruptured intracranial aneurysm (UIA) remains controversial. In Japan, medical-related disputes between patients and healthcare providers regarding decision making on UIA management has been gradually increasing with cases brought to court.

Due to the controversy, several decision analyses have been conducted and published [2-4]. A decision analytic framework could provide comprehensive information regarding what is known and unknown about the disease, options for treatment, possible outcomes, and encourage elicitation of preference for outcomes [5]. In addition, results of the analysis could provide recommendations, important factors for decision making, and a framework for discussion with a physician.

Despite the above scientific merits for physicians and researchers, however, healthcare consumers may not be able to fully utilize the evidence-based information. Difficulties include perception of uncertainty, preference measurement, and interpretation of the result in application to "my case".

In this study, we developed, implemented, and performed a preliminary evaluation of a web-based decision analytic tool for healthcare consumers called u-SHARE (ubiquitously-Support and Heal patients with intracranial Aneurysms with Risk communication and Empowerment).

Materials and methods

Conceptual framework

Figure 1 demonstrates the conceptual framework for decision making by UIA patients using a decision analytic approach.

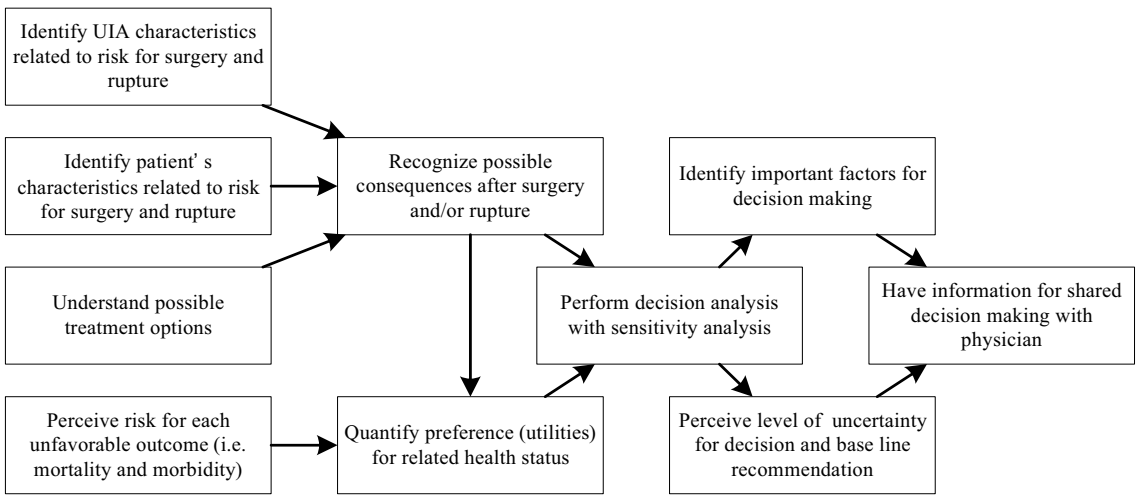


Figure 1 - Conceptual model for a decision analysis support system

Our ultimate goal is to equip patients with enough information and knowledge for appropriate shared decision making with their physicians. To achieve this goal, u-SHARE is designed to provide systematic steps for patients to understand (1) their UIA and physical conditions, which influence the risk of rupture and surgery, (2) treatment options, (3) possible consequences after surgery or rupture, and (4) their preference in unfavorable outcomes.

Decision analytic model

The published Markov decision analytic model was used in this study [2]. The decision analysis compared preventive surgery for a detected UIA and "watchful waiting", with quality adjusted life years (QALYs) as an outcome. The decision analytic model includes four health states: well, living with UIA, severe neurological deficit defined as a Rankin Scale of 3-5, and death. Quality of life with UIA and severe neurological deficits are designed to be collected from each user.

Although the original paper utilized probability data from a previous ISUIAI report [1], we updated the data with the current ISUIAI report published in 2003 [6]. The ranges for on-way sensitivity analysis for probabilities were defined with 95% confidence intervals. The ranges for utilities were set as 0 – 1.0. The decision model was developed with TreeAge Pro 2006 software (TreeAge Software, Inc., Williamstown, MA. 2006).

Table 1 reveals key probability data for the decision analysis.

Table 1 – Baseline data for decision analysis system design

Items	Baseline
Preventive Surgery	
Mortality	4.5%
Physical complication	68.6%
Relative Risk	
Age 50	2.4
Past history of strokes	1.9
Aneurysmal symptoms	1.6
Annual Risk of Rupture	
Anterior UIA: < 7 mm	0.5%
Anterior UIA: 7 - 12 mm	3.1%
Anterior UIA: 13 - 24 mm	10.2%
Aneurysmal rupture	
Mortality	43.1%
Physical complication	24.3%

Only part of the data is shown due to space limitation.

The u-SHARE application is developed as a web-based application run by Active Server Pages (ASP). Default probability data reported in the ISUIAI report [6], such as surgical mortality and morbidity for preventive surgery and annual rupture rate for various type of UIA, is stored in a relational database management system (RDBMS) developed in MySQL. The system schema of u-SHARE is illustrated in the Figure 2.

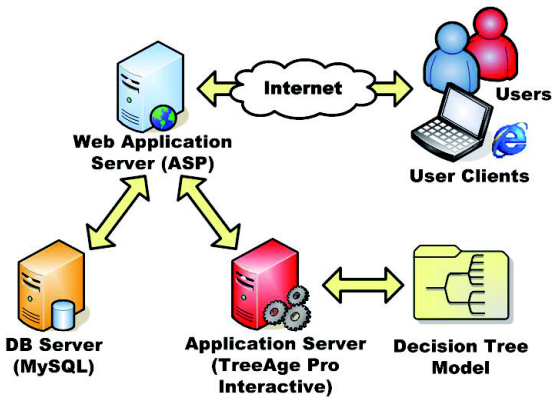


Figure 2 - Web-application overview

The web server, MySQL RDBMS, and the decision analytic model created with TreeAge Pro are actively connected to one another by a special ASP program called TreeAge Pro Interactive (TAPI) (TreeAge Software, Inc., Williamstown, MA. 2006).

All data entered through the web site is passed from the ASP server and stored in the MySQL RDBMS via ODBC (Open DataBase Connectivity). TAPI is an ActiveX interface application, which connects data from the MySQL database and the web browser to a decision model developed by the TreeAge Pro software. Active X control enables users to employ any command in the TreeAge Pro application through the web site. In addition, the results can be transferred to the TAPI which can dynamically generate an HTML file.

Preliminary evaluation

A total of 29 testers (14 healthcare professionals and 15 non-medical individuals) were asked to evaluate the u-SHARE. We modified a questionnaire, which was reported in a previous MedInfo paper, for IT decision support system evaluations [7]. The questionnaire consists of 13 questions corresponding to the following three assessment categories

- Content appropriateness - corresponding to the conceptual model described in Figure 1
- Usability
- Clinical usefulness for shared decision making

Table 2 shows all questions in the questionnaire.

Questions are asked with a scaling of 1 to 7 (1= strongly disagree, 2= disagree, 3= somewhat disagree, 4= somewhat, 5= somewhat agree, 6= agree, 7= strongly agree).

Participants were also interviewed to gain further detailed comments regarding the general concept of a web-based decision analytic tool.

Table 2 – Questions for preliminary evaluation

No	Questions
Content appropriateness	
1	Can you confirm the type of UIA you have?
2	Do you perceive the risk of unfavorable outcomes?
3	Do you understand the available treatment options?
4	Do you understand the health conditions of severe neurological deficit?
5	Was your preference for the severe neurological deficit successfully elicited?
Usability	
6	Is the text readable?
7	Are the images easy to recognize?
8	Did you have difficulty proceeding with the analysis?
9	Is appropriate guidance provided?
Clinical usefulness for shared decision making	
10	Do you understand the recommendation?
11	Do you recognize the uncertainty of decision making?
12	Do you perceive important factors for your decision?
13	Would you want to use this tool for decision making?

Results

The u-SHARE is currently available on the web site (http://www.u-share.org). The following is a description of u-SHARE.

Description of the web-based decision analytic tool

Visitors of u-SHARE will obtain general background information, current medical evidence, basic steps for decision analysis and understand the limitations of the decision analytic tool. Each UIA patient is required to register a username and password in the system so they can return to the system for further decision analysis.

The patient is asked to input factors related to risks, reported in an ISUIAI paper [6], such as age, gender, presence / absence of aneurysmal symptoms other than rupture and previous ischemic cerebrovascular disease.

In the next screen, patients are required to enter UIA characteristic data (i.e. size, location) which are closely related to risks for surgery and rupture.

Once, the UIA data is entered, graphical information about the annual rupture rate, mortality and morbidity after UIA rupture, and surgical mortality and morbidity are summarized as pie charts. (Figure 3) .

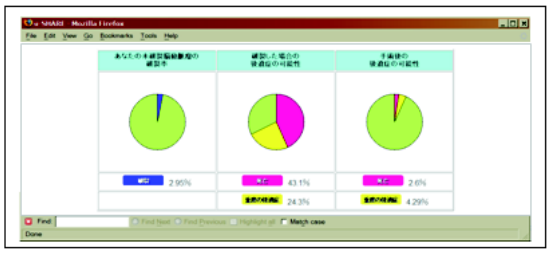


Figure 3 – Risk of Surgery and Rupture

After the patient understands the risks and types of unfavorable outcomes, u-SHARE explains the average health condition in regards to the morbidity (neurological deficit defined by a modified Rankin Scale of 3-5) [8]. The explanation is made with written sentences and a voice and video clip using Windows Media Player. Figure 4 demonstrates examples of the video clip demonstrating severe neurological deficit.

Figure 4 – Examples of the video clip

After understanding the health states, u-SHARE will elicit utility values from patients for a QALYs calculation. Figure 5 demonstrates a tool for utility elicitation using a time trade-off (TTO) method [9].

After examining the necessary information for decision analysis (i.e., probability data for each UIA and utility for health states), the patient can perform a decision analysis. In order to visualize and emphasize uncertainty of the results, we employed a Monte Carlo simulation to display the recommendation. (Figure 6).

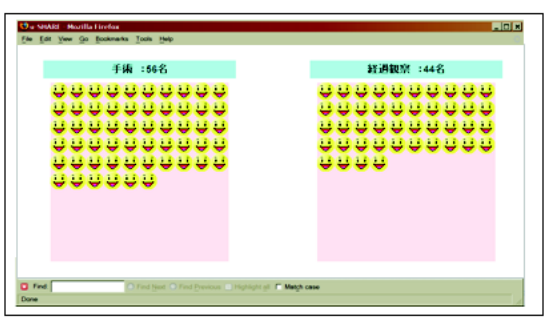


Figure 5 – Graphical results of the Monte Carlo simulation

The Monte Carlo simulation in Figure 5 reveals the proportion of patients who benefit from each strategy. For instance, Figure 6 shows 56% of UIA patients benefit from preventive surgery, and 44% benefit from watchful waiting. This format was selected as we intend to provide information for decision making rather than a number / recommendation.

u-SHARE then performs a one-way sensitivity analysis for annual UIA rupture rate, morality and morbidity after preventive surgery, mortality and morbidity due to rupture, quality of life with UIA, and assigns a Rankin Scale between grades 3-5. The factors within the predefined threshold range, appear as “important factors for decision making” in the following format. (Figure 6). The graph reveals current value, threshold, and range of probability or utility for further consideration and discussion.

Figure 6 – Graphical demonstration of threshold analysis

Finally, all of the information used in the analysis and results by u-SHARE is summarized in a printable format. We expect that each patient would use the summary for further discussion with their family, friends, or healthcare professionals.

Preliminary evaluation

Table 3 shows results of evaluation. Decimal number indicate average score (7= strongly agree to 1= strongly disagree), and percentage shows proportion of testers who put score greater than or equal to five. .

Content appropriateness

Generally, the average score was greater than 4.0, indicating that most testers felt that u-SHARE had appropriate content. Healthcare professionals scored higher than non-healthcare ones, which indicated that more detail guidance may be needed for this latter group although the contents do include appropriate information of risk and outcomes.

Ease of use

Although characters and images in the u-SHARE appear easy to understand (Qs 6 and 7), more than half non-healthcare professionals felt comfortable with the decision analysis, but more than half of the physicians felt uncomfortable about it (Qs 8 and 9).

Clinical usefulness for shared decision making

More than 85% of non-healthcare professionals successfully recognized the recommendation and its uncertainty, which was a higher proportion than for healthcare professionals. As well, more than 70% of non-healthcare professionals understood "important factors for decision making", which was also a higher score than for the healthcare professionals'.

Table 3 - Result of evaluation

No	Healthcare Professionals		Non-Healthcare individuals	
1	5.2	(71.4%)	4.9	(66.7%)
2	5.7	(85.7%)	5.3	(73.3%)
3	5.9	(85.7%)	4.4	(46.7%)
4	5.8	(92.9%)	5.2	(73.3%)
5	5.6	(78.6%)	5.7	(86.7%)
6	5.1	(71.4%)	5.5	(66.7%)
7	5.1	(64.3%)	5.2	(73.3%)
8	4.1	(35.7%)	4.6	(53.3%)
9	4.3	(42.9%)	4.9	(66.7%)
10	5.1	(71.4%)	5.3	(86.7%)
11	5.1	(64.3%)	5.3	(86.7%)
12	4.6	(50.0%)	5.0	(73.3%)
13	4.1	(42.9%)	4.4	(46.7%)

Discussion

This study demonstrated that the integration of current information technology could overcome the difficulty of deploying decision analysis systems. Since a large barrier to utilizing decision analysis is requirement of special knowledge, even for clinicians, this web-based tool could be an alternative format of publication for practical usage of decision analysis in clinical settings in addition to scientific reports in medical journals.

In addition, the study showed decision analysis systems were accepted as an evidence-based second opinion acquisition source for autonomic decision making by patients. Our preliminary evaluation implied that decision analytic tool might be more acceptable for healthcare consumers than for healthcare professionals. Such a tool may help healthcare consumers to make appropriate decisions, and as such could be considered as a form of structured informed consent

Finally, during development of u-SHARE the authors recognized the difficulty in developing common understanding among professionals in healthcare, information technology specialists, and decision analysts. It is important to encourage collaboration among people who can produce interdisciplinary research and development which combine various disciplines to generate marketable, medically significant, and sustainable services, especially in the health informatics area.

Conclusion

Our study showed that a web-based decision analytic tool could have significant potential in assisting healthcare consumers achieve appropriate shared decision making and which can also be used as "structured informed consent". A clinical trial is planned, as the logical next step for the evaluation, with actual UIA patients.

Acknowledgments

The authors thank the CAC, Inc. (Tokyo, Japan) for providing programming skills.

References

[1] International Study of Unruptured Intracranial Aneurysms Investigators. Unruptured intracranial aneurysms--risk of rupture and risks of surgical intervention. N Engl J Med 1998: 339(24): 1725-1733.

[2] Aoki N, Beck JR, Kitahara T, et al. Reanalysis of unruptured intracranial aneurysm management: Med Decis Making 2001: 21(2): 87-96.

[3] Aoki N, Kitahara T, Fukui T, et al. Management of unruptured intracranial aneurysm in Japan Med Decis Making 1998: 18(4): 357-364.

[4] Dippel DW, Vermeulen M, Braakman R, et al. Transient ischemic attacks, carotid stenosis, and an incidental intracranial aneurysm. A decision analysis. Neurosurgery 1994: 34(3): 449-457; discussion 457-448.

[5] Sanders GD, Hagerty CG, Sonnenberg FA, et al. Distributed decision support using a web-based interface: prevention of sudden cardiac death Med Decis Making 1999: 19(2): 157-166.

[6] Wiebers DO, Whisnant JP, Huston J, 3rd, et al. Unruptured intracranial aneurysms: natural history, clinical outcome, and risks of surgical and endovascular treatment Lancet 2003: 362(9378): 103-110.

[7] Aoki N, Ohta S, Masuda H, et al. Edutainment tools for initial education of type-1 diabetes mellitus - initial diabetes education with fun Medinfo 2004: 2004: 855-859.

[8] Bonita R, and Beaglehole R. Recovery of motor function after stroke Stroke 1988: 19(12): 1497-1500.

[9] Torrance GW. Utility approach to measuring health-related quality of life J Chronic Dis 1987: 40: 593–600.

Address for correspondence

Noriaki Aoki, MD, PhD, MS, FJSIM, CPE, Assistant Professor, School of Health Information Sciences, University of Texas Health Science Center at Houston, 7000 Fannin, UCT-600, Houston, TX 77030, USA;
TEL: (713) 500-3938
Email: Noriaki.Aoki@uth.tmc.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Health On the Net Foundation: Assessing the Quality of Health Web Pages All Over the World

Célia Boyer[a], Arnaud Gaudinat[a], Vincent Baujard[a], Antoine Geissbühler[a,b]

[a]*Health on the Net Foundation, Geneva 14, Switzerland*
[b]*Service of Medical Informatics, Geneva University Hospitals, Switzerland*

Abstract

The Internet provides a great amount of information and has become one of the communication media which is most widely used [1]. However, the problem is no longer finding information but assessing the credibility of the publishers as well as the relevance and accuracy of the documents retrieved from the web. This problem is particularly relevant in the medical area which has a direct impact on the well-being of citizens. In this paper, we assume that the quality of web pages can be controlled, even when a huge amount of documents has to be reviewed. But this must be supported by both specific automatic tools and human expertise. In this context, we present various initiatives of the Health on the Net Foundation informing the citizens about the reliability of the medical content on the web.

Keywords:

patient safety, natural language processing, internet, information retrieval, information quality

Introduction

The Internet has become one of the communication media that is most widelyused. With the availability of Web server software, anyone can set up a Web site and publish any kind of data which is then accessible to everyone. The problem is therefore no longer finding information but assessing the credibility of the publishers, as well as the relevance and accuracy of the documents retrieved from the web. This problem is particularly relevant in the medical area: the information found (explanations of diseases, recommended treatments, available medicine ...) has a direct impact on the well-being of the citizens. Indeed, in many cases, web sites provide no appropriate information regarding the scientific design of a medical study, nor are studies made available that support given claims. Various initiatives have been proposed for managing the quality of web pages:

- *Selection or referencing* of pages was first implemented by the *Yahoo!*[1] directory. In the medical area, the American *MedlinePlus* [2] and the French *CISMeF* [3] portals are referred to. The principle of referencing consists of publishing a selection of web pages arranged according to domains which are organised more or less hierarchically.
- *The accreditation* of web pages, such as performed by HONcode [4] and URAC [5], specifically aims at promoting the quality of information. According to this intiative, sites must respect a set of criteria to be accredited.
- *The popularity* of web pages is detected by the use of PageRank [6] or similar models. The popularity is correlated with the judgement of webmasters when they decide whether links to other web pages should be estalblished.
- *The social or collaborative initiative* consists of using the judgment of all users in order to characterise web content and thus, to create networks of reliable users in each addressed area. *Lijit* [7] and *Google co-op* [8] implement this initiative.
- *The self-regulation* of web sites consists of their auto-accreditation by respecting given rules [9,10].
- *The education of patients* consists of helping them to find reliable web sites on the basis of a set of criteria [11,12].

The objective of the Health on the Net Foundation (HON[2]) is to provide citizens with information on the reliability of the medical content of web documents. HON assumes that the quality of web sites should be certified by a neutral third party organisation such as HON, according to quality criteria such as the HONcode. Since 1995, the HONcode has been accepted by the health web community. In 2001, HON Foundation was recognized as a Non Governmental Organisation and currently provides consulting services to ECOSOC. Moreover, HON has participated in the workgroup of the European Community being in charge of the elaboration of eEurope 2002: Quality Criteria for Health related Websites. And HON is a partner of WHO for the development of quality health information in African countries. In addition to the experience acquired by the HON Foundation, we believe that the difficulty of promoting high quality health information on the web, can be reached by combining the development of suitable automatic tools with human expertise. In the rest of this paper, we introduce the HON Foundation by presenting three

1 www.yahoo.com

2 Health On Net: *http://www.healthonnet.org*

major aspects of its activity: (1) the HONcode of Ethical Conduct, (2) the database of accredited medical web sites and (3) the automatic tools created in order to help users and human reviewers when assessing the quality of medical sites. We then draw a conclusion and provide some perspectives to the HON Foundation activity.

HONcode principles and accreditation protocol

Ethical HONcode of conduct

The HONcode [4] is a set of high ethical principles defined by the HON Foundation in order to assess the quality of on-line health information on the web. These principles are applied during an evaluation of the web sites which have demanded to be accredited by HON. Each principle has both, short patient-oriented (here above) and long expert-oriented formulations:

1. *Authoritative*: Indicate the qualifications of the authors
2. *Complementarity*: Information should support, not replace, the doctor-patient relationship
3. *Privacy*: Respect the privacy and confidentiality of personal data submitted to the site by the visitor
4. *Attribution*: Cite the source(s) of published information, date and medical and health pages
5. *Justifiability*: Site must support claims relating to benefits and performance
6. *Transparency*: Accessible presentation, identities of editor and webmaster, accurate email contact
7. *Financial disclosure*: Identify funding sources
8. *Sponsorship*: Clearly distinguish advertising from editorial content

The long description of principles is more explicit and should remove any remaining uncertainties about their meaning. For instance, the complete statement of the first principle *Authoritative* is: *Any medical or health advice provided and hosted on this site will only be given by medically trained and qualified professionals unless a clear statement is made that a piece of advice offered is from a non-medically qualified individual or organisation.* It indicates particularly the interest for medical information being provided by medically trained people. The HONcode principles can be read online in 32 languages. Additionally, The HON Foundation provides the guidelines of the HONcode accreditation process [13].

Accreditation process and active seal

During the accreditation process, human experts evaluate web sites, which have demanded for the accreditation, in order verify if these sites respect the ethical HONcode. If web sites respect the principles, the dated accreditation seal (Figure 1) and the personal identification number (PIN) corresponding to the accreditation certificate are provided to these sites. The provided seal is *active*, and when one clicks on it in order to verify its validity, the generated screen contains the personalised HON page presenting the Web site's accreditation status. This technique allows us to reflect in the status of an accredited web site in real time and, if needed, to change it from "accredited" to "under reviewing process" or "not compliant" (see Figure 1).

Figure 1 - Active Seals, "accredited", "under review" and "not compliant" from the certificate

During the accreditation review, the HON experts identify modifications that are required for the site in order to be accredited. Each missing ethical information is indicated and must be added to the content of web site pages. Once this has been done, the accreditation seal and PIN are provided. All the accredited web sites are reviewed each year. If a web site does not anymore respect the HONcode principles, the webmaster receives warnings and, if required modifications are not made, the site loses its accreditation. The HONcode ethical principles and accreditation are being adopted by health web publishers all over the world. The accreditation process, as it is implemented, allows to conduct an educational work with webmasters and to naturally improve the quality of on-line health information. The accreditation request can be made on-line. A pre-evaluation is proposed to the webmaster in order to highlight the missing principles.

Accredited HONcode database of websites

Currently, the database is composed of over 5,500 accredited web sites in 72 countries which represents over 1,200,000 web pages indexed in Google. 52% of the accredited sites are in English and about 11% of the sites in French, followed by a considerable amount in Spanish, Italian, Portuguese and Dutch. For each evaluated evaluated site the database provides the following information, which assists us in developing automatic tools:

- respect of ethical HONcode labels,
- excerpts corresponding to these eight principles,
- indexing with MeSH headings (i.e., *diabetes*, *asthma*),
- general content type labels (i.e., *Woman health*, *Information for Patients*).

Automatic tools

The automatic tools developed by the research and development team of the HON Foundation address two main objectives:

- help citizens to manage the ever increasing amount of information on the web and provide them with reliable medical web resources,
- help human reviewers during the web sites' evaluation and reviewing processes.

According to these objectives, we addressed the following tasks when conceiving and developing tools: crawling task with MARVIN tool, indexing of medical pages with the MeSH terms, selection of medical content, development of

the specific HON toolbar, and automatic detection of statements on HONcode principles. The majority of these tools can be tested on-line through the HON Foundation web site (http://www.healthonnet.org). They can be used by medical experts (doctors, nurses, students) and citizens.

MARVIN the crawler

The objective of the MARVIN (Multi-Agent Retrieval Vagabond on Information Networks) crawler [14] is not only to find web pages which respond to the query, but to locate the right piece of information among them. Indeed, general web search engines, indexing most of the web, return a long list of documents, often to the detriment of precision. The search result is then barely usable because of the large number of answers from different domains and topics. Only complex queries may, in a given situation, produce a limited number of potentially relevant documents. In order to make searches more efficient and useful to ordinary users, intelligent and specialised search engines are needed on the web. The HON crawler MARVIN was first applied to the medical domain. Armed with a dictionary of medical terms and the MeSH terminology, MARVIN skims the Web for new sources of medical information.

Automatic indexing of web pages with the MeSH keywords

MeSH [15], Medical Subject Heading is a terminology developed by the National Library of Medicine. It contains currently over 33,000 terms. MeSH is translated into various languages. In our applications, we use its versions in English, French, German, Spanish, Portuguese, Danish, Dutch and Italian. MeSH has been conceived for information retrieval purposes. Its hierarchical organization thus offers important search opportunities by using the "concept exploding". For instance, a search with *diabetes mellitus* will also bring documents indexed with the following terms: *diabetes mellitus, experimental, diabetes mellitus, type 1, diabetes mellitus, type 2, diabetes, gestational, diabetic ketoacidosis,* and *prediabetic state*. Indeed, all these terms are hierarchical children of the term *diabetes mellitus*, and the concept exploding function will "explode" *diabetes mellitus* to all its children and, possibly, to other related terms. Through the interfaces of our search engines, the user may select the terms of interest and refine his or her search. The multilingual capability of MeSH transforms the MeSH thesaurus into a search tool which is even more powerful. Several of our tools exploit the multilinguality of the MeSH.

In our approach the indexing of web pages with MeSH terms is performed in two steps there are: the extraction of these terms and their weighting. The extraction of the MeSH terms relies on lexical normalisations and the detection of synonymous relations between terms. The weighting of terms has first been performed by learning algorithms, which showed interesting results. Furthermore, in order to improve these results, we exploited semantic relations between terms, as recorded in the table of co-occurring terms of the UMLS [16]. The UMLS (Unified Medical Language System) is a knowledge source of the biomedical area in which various terminological resources have been merged. The UMLS provides, among other things, the file of co-occurrences that is computed on all the MEDLINE databases. In order to compute the weighting of terms on the basis of these co-occurrences, we assume that the more co-occurring a concept is with other concepts, the more important is this concept. During the reranking process of extracted terms according to this methodology, we proceed in two steps [17]:

- for each pair of terms in the list of extracted terms, we calculate the cumulative weight of their relations;
- we express each term according to its related pairs to obtain a unique score by the term.

Evaluation of this indexing system in comparison with similar French systems (CISMeF-NLP [18] and NomIndex [19]), showed that the HON MeSH-indexer presents the best F-measure [20] as compared with manually built gold standard.

Hunt and HONselect services

Hunt services are based on a flexible search engine especially tailored to specific domains. Each Hunt database can further be organised according to a specific domain such as the medical one or the domain medical biology. For instance, the MedHunt [21] database is dedicated to the health domain and is composed of documents which are automatically retrieved from the web and selected by the HON team. In addition MedHunt is capable of narrowing down a search to Web sites in either French or English, and it has currently also been updated to function in German.

HONselect [22] is a multilingual search tool integrating heterogeneous web resources. It covers the languages English, Dutch, French, German, Italian, Spanish, and Portuguese. HONselect offers a comprehensive collection of medical terms and corresponding pictures, bibliographic references, news and Web sites. To our knowledge, no other search engine offers such easiness in searching the medical web. HONselect integrates several separate databases (Medline, HONmedia, NewsPage, MedHunt, HONcode, clinical trials as well as scientific articles). Moreover, the HONselect proposes more options: (1) its *search* function checks automatically for spelling errors in seven languages; (2) its *translate* function allows to easily switch from one language to another so one can obtain additional and complementary information in other languages. Search results are displayed in the language of the formulated query.

Currently, the crawling task is also supported by the Google Co-op Health service. Indeed, it becomes possible to define one's own search engine on the basis of the ommon Google crawler. According to the HON Foundation policy, the tailoring of the Google crawler is performed by providing a list of trustworthy web sites from the HONcode accredited database. This helps to increase the precision of obtained results. In the mean time, the performance and coverage of the Google crawler and its storage capacity helps to increase the recall. Com-

bining expertise of both the HON Foundation and the Google incorporation, we can contribute to satisfy the citizens when they look for reliable health information on the web.

Verification of web pages' reliability with WRAPIN

The WRAPIN engine [17] (Worldwide online Reliable Advice to Patients and Individuals) has the objective to enhance the capabilities of HON's retriever and indexer MARVIN. The innovations gained by the WRAPIN engine, in addition to the already integrated functions are:

- processing of health and medical documents in any format (HTML, PDF, etc) or length,
- querying of more medical and health databases: i.e., Bookshelf, ClinicalTrials, PubmedCentral, PubMed, FDA, OESO and Urofrance,
- determining information quality by comparing the documents with the interconnected knowledge base,
- providing a summary of the ideas contained in the processed documents,
- using of the entire document and of the URL as the search query, in addition to a normal series of search terms,
- highlighting of MeSH terms in the summary in order to better define the correctness of results,
- identifying results according to the user's profile (specialist or new to the field of medicine).

Automatic identification of HONcode accredited sites

For the automatic identification of HONcode accredited web sites, The HON Foundation proposes a toolbar which detects the HONcode status of a web site. This toolbar has been accomplished within the framework of the European project ActiveHealth (Active Environment for Health Promotion and Disease Prevention). The aim of this project was to transfer information to individuals. This toolbar is downloadable from the HON web site [23]. It has two functions:

- Automatic checking of the accreditation of the web site being read by the user;
- Searching in the HONcode accredited web sites: (1) by entering a word or phrase into the search box or (2) by selecting text using the mouse, and right-clicking to begin the search.

Automatic quality criteria extractor

The proposed methodology for the automatic extraction of quality criteria focuses on processing vast amounts of health data on the Internet. An automatic tool [24] is currently under development in order to help the daily work of human reviewers when they evaluate the transparency of health websites and annotate them by using the HONcode principles. This tool relies on the application of machine-learning algorithms such as Naive Bayes, Support Vector Machine, k-Nearest Neighbours and Decision Tree [25,26]. This tool is trained on the HON database of accredited web sites. It is currently applied to medical web sites and shows up to 78% of precision and 73% of recall. The contingency computed between precision and recall indicates that some of the principles of the HONcode are rather easy to detect. For instance, the *Privacy* principle shows one of the best contingency rates, 92% / 90%. While other principles remain problematic: the *Justifiability* shows the precision/recall contingency of only 45% / 33%. Moreover, it appears to be ambiguous with the *Complementarity* principle.

Conclusion and perspectives

During the past ten years, the HON Foundation has undertaken important efforts in order to promote the quality of health information on the web worldwide. To address the size of the web, the necessary human expertise is supported by various automatic tools (crawler, indexer, search engine, HONcode principle extractor) in order to make the evaluation of web sites more systematic and rapid. From the user's point of view, easy mechanisms are provided such as the specific toolbar and the dynamic seals, and particularly, a database of high quality medical web sites.

All the methodologies and tools developed by the HON Foundation can also be applied to other areas, where trust and transparency are important issues.

Additional solutions can be found in order to assist the citizens and medical professionals even more in searching the medical web. Some of these conceptions can be realised with partners from the addressed areas, such as the National Library of Medicine, Bethesda, USA or Google Inc.

By proposing the HONcode, HON educates information providers to be more transparent. However, users are not aware enough regarding the coexistence of reliable and unreliable information on the web. HON should conduct awareness campaigns in order to educate the citizens to efficiently use the medical and health information on the Web. Collaboration and integration with search engine mainly used is also crucial to directly reach the user in his/her daily life. So, HON has worked closely with Google to highlight the trustworthy health information to the citizen by using a set of health labels they jointly developed.

The approach so far conducted by HON is a global one covering 32 languages worldwide. However, HON should respond to the local needs and the various languages spoken. The creation of chapters located in different regions of the world might help us to locally and thus globally improve the quality of medical and health information.

Acknowledgments

The foundation activities are supported by the State of Geneva and more particularly by the Geneva Ministry of Health (DES) and the European Union through the WRAPIN, ACTIVEHealth and PIPS projects.

References

[1] Fox S. Online Health Search 2006, Online Health Search 2006. Washington, DC: Pew Internet & American Life Project, October 29, 2006.

[2] Miller N, Lacroix EM and Backus JE. MEDLINEplus: building and maintaining the National Library of Medicine's

consumer health Web service. *Bull Med Libr Assoc* 2000;88(1): 11-7.
[3] Darmoni SJ, Leroy JP, Baudic F, Douyere M, Piot J and Thirion B. CISMeF: cataloque and index of French speaking health resources. *Stud Health Technol Inform* 1999:493-6.
[4] Boyer C, Baujard V and Scherrer J. HONcode: a standard to improve the quality of medical/health information on the internet and HON's 5th survey on the use of internet for medical and health purposes. In *6th Internet World Congress for Biomedical Sciences (INABIS 2000)*, 1999.
[5] *www.urac.org/MMandQualityChasm.asp*
[6] Page L, Brin S, Motwani R and Winograd T. The pagerank citation ranking: Bringing order to the web. *Technical report*, Stanford Digital Library Technologies Project, 1998.
[7] *www.lijit.com*
[8] *www.google.com/coop*
[9] Risk A. Dzenowagis J. Review of Internet information quality initiatives. *Journal of Medical Internet Research.* 2001; 3(4):e28
[10] Sheldon T. Trust mark launched as a guarantee of safety in the Netherlands. *BMJ* 2002;324(7337):567
[11] Charnock D. The DISCERN Handbook. 1998. Radcliffe Medical Press
[12] *omni.ac.uk*
[13] *www.hon.ch/HONcode/Guidelines/guidelines.html*
[14] Baujard O, Baujard V, Aurel S, Boyer C and Appel R. A multi-agent softbot to retrieve medical information on internet. *Medinfo* 1998; 150–4.
[15] National Library of Medicine, Bethesda, Maryland. Medical Subject Headings, 2001. http://www.nlm.nih.gov/mesh/meshhome.html.
[16] NLM. UMLS Knowledge Sources Manual. National Library of Medicine, Bethesda, Maryland, 2005. www.nlm.nih.gov/research/umls/.
[17] Gaudinat A, Joubert M, Aymard S, Falco L, Boyer C, and Fieschi M. WRAPIN: new generation health search engine using UMLS knowledge sources for MeSH term extraction from health documentation. In M. F. at al., editor, *Medinfo*, p. 356–360, 2004.
[18] Névéol A, Rogozan A and Darmoni S. Automatic indexing of online health resources for a french quality controlled gateway. *Information Processing & Management*, 42(3):695–709, 2005.
[19] Pouliquen B, Delamarre D and Le Beux P. Indexation de textes médicaux par extraction de concepts, et ses utilisations. In *JADT*, pages 617–628, 2002.
[20] Névéol A, Mary V, Gaudinat A, Boyer C, Rogozan A and Darmoni SJ. A benchmark evaluation of the french mesh indexing systems. In S. Miksh, J. Hunter, and E. Keravnou, editors, *Artificial Intelligence in Medicine*, Springer's Lecture Notes in Computer Science, pages 251–255, 2005.
[21] Boyer C, Baujard O, Baujard V and Appel R. New developments of MedHunt medical search engine. In *MedNet 98*, Brighton, U.K., 1998.
[22] Boyer C, Baujard V, Nater T, Appel R and Scherrer J. Honselect: A new assisted-search facility for health information and resources integrating heterogeneous databases. In *MedNet 99*, Heidelberg, Germany, 1999.
[23] *www.hon.ch/HONcode/Plugin/Plugins.html*
[24] Gaudinat A, Grabar N and Boyer C. Utilisation de méthodes d'apprentissage pour la détection des critères de qualité des documents en ligne du web médical. In O. Ly and C. O. Bayaoko, editors, *JFIM 2007*, Bamako, Mali, 2007. To appear.
[25] Grobelnik M and Mladenic D. Efficient text categorization. In Mining workshop on the *10th European Conference on Machine Learning ECML98*, 1998. citeseer.ist.psu.edu/grobelnik98efficient.html.
[26] Manevitz LM and Yousef M. One-class SVMs for document classification. *Journal of Machine Learning Research*, 2:139–154, 2001.

Address for correspondence

Célia Boyer, Health on the Net Foundation, HUG/SIM
24, rue Micheli-du-Crest, 1211 Geneva 14 Switzerland
tel: +41 22 372 62 50 fax: +41 22 372 88 85
email: celia.boyer@healthonnet.org

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

The Use of Electronic Medication Reconciliation to Establish the Predictors of Validity of Computerized Medication Records

Alexander Turchin, MD, MS[a,b,d], Tejal K. Gandhi, MD, MPH[b,d], Christopher M. Coley, MD[c,d], Maria Shubina, DSc[b], Carol Broverman, PhD[a]

[a] *Clinical Informatics Research and Development, Partners HealthCare, Boston, MA*
[b] *Brigham and Women's Hospital, Boston, MA*
[c] *Massachusetts General Hospital, Boston, MA*
[d] *Harvard Medical School, Boston, MA*

Abstract

Medication records in clinical information systems (CIS) are frequently inaccurate, leading to potentially incorrect clinical decisions and preventing valid decision support interventions. It is not known what characteristics of electronic medication records are predictive of their validity.

We studied a dataset of 136,351 electronic medication records of patients admitted to two academic hospitals that were individually validated by admitting providers using novel medication reconciliation software. We analyzed the relationship between characteristics of individual medication records and the probability of record validation using a multivariable linear regression model.

Electronic medication records were less likely to be validated if more time had passed since their last update (14.6% for every 6 months), if they represented an antiinfective (61.6%) or a prn (50.9%) medication, or if they were in an outpatient CIS rather than on an inpatient discharge medication list (18.1%); $p < 0.0001$ for all.

Several characteristics of electronic medication records are strongly associated with their validity. These findings could be incorporated in the design of CIS software to alert providers to medication records less likely to be accurate.

Keywords:

clinical information systems, electronic medical records, medications

Introduction

A growing body of literature supports the overall superior safety of clinical information systems that capture medication orders [1]. Accurate medication information in clinical information systems is therefore crucial for patient safety and quality of care [2]. Nevertheless, many investigators report that electronic medication data is frequently incomplete and / or outdated [2-4]. Other sources of medication information (e.g. patients, insurance claims or pharmacies) can be helpful but may also be inadequate [5, 6]. Consequently providers frequently face the task of identifying inaccurate medication information in the clinical information systems in absence of other clues. However, it is not known what characteristics of medication records in clinical information systems are predictive of the validity of the records.

Since January 2006, the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) has mandated that medication information be reconciled across the continuum of care [7], with a particular emphasis on transitions of care, such as hospitalizations. At our integrated healthcare delivery network we have developed a software application, Pre-Admission Medication List (PAML) Builder, that is used to document the complete list of medications the patient was taking prior to the admission to the hospital in a structured electronic format using a combination of information transfer from existing clinical information systems, electronic lists of discharge medications from previous hospital admissions and new data entry (described in more detail below) [8]. During the process of medication reconciliation using the PAML Builder, providers admitting the patient to the hospital validate or discard each of the existing electronic medication records based on the information obtained from the patient / caregiver interview and other relevant sources, and the results of these activities are permanently stored in our data warehouse. We have therefore analyzed the data generated during the medication reconciliation process to identify the characteristics of medication records that are predictive of their validity.

Materials and methods

PAML builder and medication reconciliation

The PAML Builder is a standalone software application designed to assist the process of medication reconciliation for hospitalized patients (Figure 1). The left side of the application presents a list of the patient's medications collected from four clinical information systems in use at Partners HealthCare: the Longitudinal Medical Record (LMR; outpatient), OnCall (outpatient), Brigham Integrated Computer System (BICS; inpatient) and Massachusetts

General Hospital Order Entry (MGH OE; inpatient). Medication records imported from the outpatient clinical information systems are all of the patient's active medications in that system, and medication records imported from inpatient systems are the last set of discharge medications from the corresponding hospital (from hospital admissions within the last 18 months). For each medication record, the PAML Builder displays the source clinical information system and the date of the last update of this medication record in the system. In the outpatient clinical information systems the date of the last update of a medication record is set to the current date when either a prescription is printed from the record or any of the data elements in the record are changed. In the inpatient systems the date of the last update is the date of the patient's discharge from the hospital. The right side of the application contains the list of the medications the patient was taking at the time of admission (PAML), as collected and validated by the admitting provider. The application allows providers to select all accurate electronic medication records from the left side for inclusion in the PAML. Patient's medications not found in the clinical information systems are then entered manually into the PAML. Creation of an accurate PAML is mandatory for all patients who are admitted to one of the two hospitals where the application has been rolled out.

Study patients and settings

We conducted a retrospective cohort study of all patients who were admitted to the Brigham and Women's Hospital and Massachusetts General Hospital between Aug ust 1, 2006 and October 30, 2006 and had a PAML created. For each patient, we obtained the PAML, the list of active medications in four clinical information systems available at the time of the creation of the PAML (as displayed on the left side of the PAML Builder application), and identified all medication records that were copied over to the PAML (i.e. validated to be accurate). The institutional review board at Partners HealthCare System approved the study, and the need for written informed consent was waived.

Measurements

A unique electronic medication record in one of the clinical information systems was the unit of analysis in this study. Only medication records active at the time of the hospital admission were analyzed. We also defined a unique medication entry as a combination of medication name, route and a particular admission; there could be more than one unique medication record corresponding to a unique medication entry (e.g. records documenting the same medication in different clinical information systems). Medication records active at the time of more than one hospital admission over the study period were analyzed separately for each admission. For every electronic medication record we identified the following characteristics:

1. *Inpatient vs. Outpatient*. Medication records that originated in the inpatient clinical information systems (BICS, MGH OE) were labeled "Inpatient".
2. *Prn (as needed) medication* (Yes or No). This category included medications prescribed to be taken as needed rather than regularly.
3. *Antiinfective* (Yes or No). This category included antibacterial, antiviral and antifungal agents administered systemically.
4. *Medication record age* was calculated as the number of months between the date of the last update of the medication record and the date of the patient's admission to the hospital.
5. *Validated*. Medication records that were copied by providers from the left side of the PAML Builder to the PAML were considered validated. If the medication's dose or frequency were changed after it was copied, it was still considered validated. If the medication route was changed, it was not considered validated because the same chemical substance administered by different

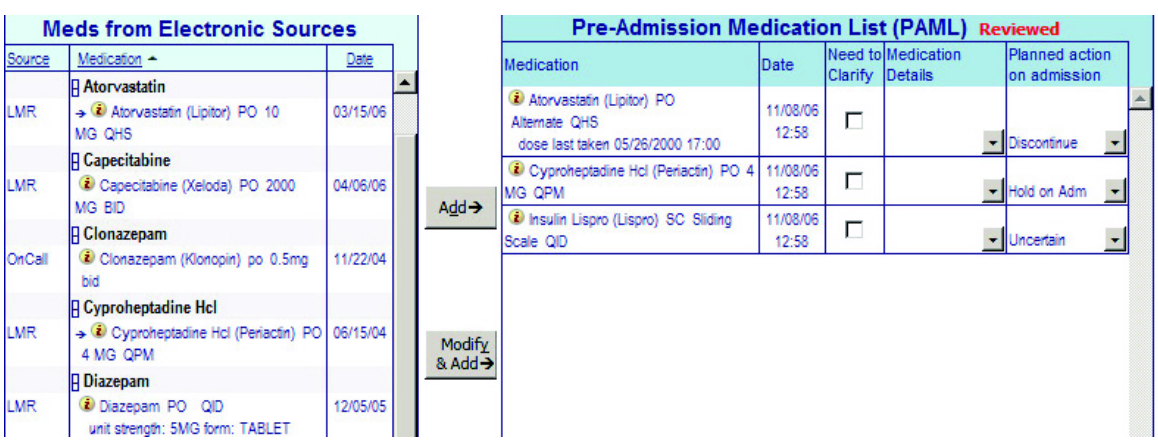


Figure 1 - Pre-Admission Medication List (PAML) Builder

The left side (Meds from Electronic Sources) displays the list of currently active medications imported from four EMRs. The "Date" column lists the date of the most recent update of the medication record in the source EMR. The right side (PAML) contains the list of medications the patient was taking prior to admission as collected and verified by the admitting provider. Users can transfer valid EMR medication records to the PAML by clicking on "Add" or "Modify and Add" buttons.

routes (e.g. oral vs. topical) typically has different indications and therefore would more likely represent a new prescription than a validation of the old one. If the medication was removed from the PAML after it was originally copied to it, it was not considered validated.

6. *Duplicate*. If there was more than one unique medication record for a given medication entry, all of the medication records corresponding to that medication entry were considered duplicates.

All data was obtained from the electronic medical records systems at Partners Healthcare.

Statistical analysis

Summary statistics were constructed by using frequencies and proportions for categorical data and by using means, standard deviations, medians, and ranges for continuous variables. We constructed a multivariable logistic linear regression model to evaluate the association between characteristics of the medication record in a clinical information system and the probability of the record being copied to the PAML (validated). All analyses were performed with SAS statistical software, version 9.1 (SAS Institute, Cary, North Carolina).

Results

Electronic medication records

We identified 17,335 hospital admissions over the study period for which a PAML was created. 10,785 (62.2%) of these admissions had at least one medication record in any of the clinical information systems and were included in the study. There were 136,351 unique medication records in all four clinical information systems linked to these admissions. Among these, there were 111,231 unique medication entries. 92,403 (83.1%) of medication entries had only one medication record linked to each of them, while 18,828 (16.9%) medication entries were linked to multiple medication records. 52.2% of study admissions had medication records only in one clinical information system, 41.6% in two systems, and 6.2% in more than two systems. On average each study admission had 12.7 medication records in the clinical information systems.

Table 1 - Characteristics of electronic medication records

Variable	Value
All records	136,351
Inpatient discharge medications	53,689 (39.4%)
PRN	21,481 (15.8%)
Antiinfectives	8,889 (6.5%)
Medication record age, months	9.8 (± 21.7)
Validated	48,286 (35.4%)

Data are means (± SD) or n (%).

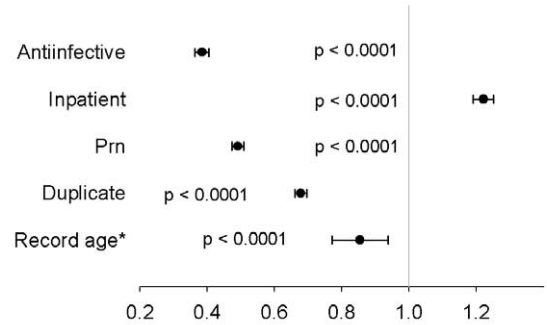


Figure 2 - Predictors of medication record validation

Circles indicate the estimated adjusted odds ratios for the medication record to be copied to the PAML (validated). Wisps indicate the 95% Wald confidence limits for the odds ratios.

* for every 6 months prior to the date of hospital admission

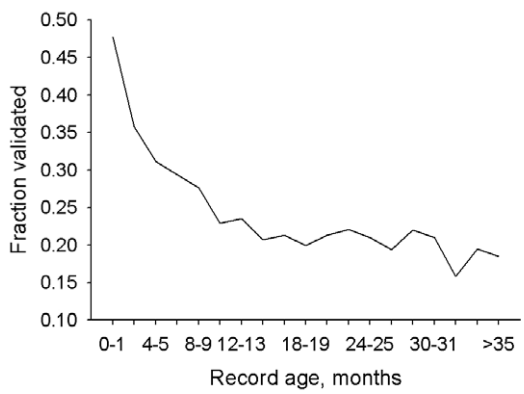


Figure 3 - Medication record age and validation

Characteristics of the medication records are shown in Table 1. Median age of the medication record was 3.5 months. 4,695 (43.5%) of study admissions had both inpatient (hospital discharge medications) and outpatient medication records, while 5,124 (47.5%) had only outpatient and 966 (9.0%) had only inpatient medication records. Over a third of medication records from clinical information systems were copied to the PAML.

Predictors of medication record validation

The effects of the characteristics of a medication record on the probability that it would be validated are illustrated in Figure 2. Inpatient discharge medications had 22.1% higher odds of being validated than the ones from outpatient EMRs. Medications ordered *prn* (as needed) and antiinfective medications had 50.9% and 61.6% smaller odds of being validated, respectively. Medications for which there were duplicate electronic records had 32.2% smaller odds of being validated. Finally, in the linear

model the odds of the medication record being validated decreased by 14.6% for every six months between the last date when the medication record was updated or the date of admission (all relationships significant at $p < 0.0001$). When the medication record age was plotted directly against the fraction of the records validated (Figure 3), it appeared that the continuous decrease in the probability of validation described by the linear model only occurred when the interval between the last update of the record and date of admission was less than one year. After one year the probability of validation stabilized at about 20%.

Discussion

In this retrospective analysis of over 130,000 medication records we have identified several characteristics of the medication records that were associated with a higher probability of record validation. As in several previously published reports [2-4], a large fraction of the medications from clinical information systems was not validated. While some investigators have reported lower prevalence of outdated medication information in clinical information systems than what we observed [9], accuracy likely varies with local practice, and based on the literature our findings appear typical.

As expected, the medication record age was an important factor and older medication records were less likely to be validated. However, further univariate analysis identified a threshold (approximately one year) after which medication record age no longer appeared to play a significant role. Only one in five medication records over one year old were validated compared to nearly half of the medication records that were updated or had a prescription printed from them within the last month. This finding confirms previous reports that a large fraction of medication records in clinical information systems may be outdated [2], and demonstrates that the date when the record was last updated or had the last prescription printed from it is an important indicator of accuracy. It corroborates previously published smaller investigations limited to a single pharmacological domain that reported that a record of the prescription being filled within the last six months had a high positive predictive value for identification of active medications [10]. As reported in other studies, medications that the patient no longer takes but that are still listed as active in the clinical information system put the patient at particularly high risk for adverse drug events [2]. It may therefore be helpful if the date of the last update or prescription for each medication record is prominently displayed in the clinical information system in order to alert providers to possible inaccuracy of an old record.

Medication records representing antiinfective agents were significantly less likely to be validated. Previous reports that focused on medications missing from clinical information systems did not find that antibiotics were missing more frequently than average [3] while other studies, similarly to ours, found that outdated antibiotic records were common [2]. It therefore appears that there is an asymmetric lack of quality of antiinfective medication data in clinical information systems. This phenomenon could likely be explained by the fact that antiinfective agents are more commonly than other medications taken only for a limited period of time. If the medication records are not set to be automatically marked as inactive after the patient has completed the course of treatment, they will remain on the patient's record until manually removed, making them susceptible to errors of omission.

Medications prescribed on as-needed basis were also less likely to be validated. Similarly to the antiinfective agents, this discrepancy was asymmetric and other studies did not report that as-needed medications were missing from the medical record more frequently than medications taken on a regular basis [11]. It is therefore likely that the explanation for the high rate of outdated medication records in this group also lies in the increased prevalence of transient prescriptions that are not subsequently manually deactivated. Based on these findings, it may be helpful to incorporate special handling of antiinfective and *prn* medications in clinical information systems design to include additional alerts for providers to inactivate these medications when the original prescription has expired.

Medication records representing discharge medications from previous admissions were slightly more likely to be validated that the ones from outpatient clinical information systems. One possible reason for this is that patients admitted to the hospital usually have their medications reconciled [7]. While this process may not always be complete [12], it could nevertheless be more comprehensive than outpatient medication reconciliation due to more extensive resources available in the hospital.

Our study has multiple strengths. It was a large-scale investigation that included over 130,000 medication records from more than 10,000 admissions at two hospitals. To our knowledge, it is the first investigation of the characteristics associated with medication record validation. We took advantage of a novel medication reconciliation software – PAML Builder – to assemble a unparalleled database of medication records that were individually validated by admitting providers. These records included patients being admitted to all hospital services, creating a clinically diverse study population. Finally, most previously published studies focused on the validity of medication records in outpatient clinical information systems but our analysis also included hospital discharge medication lists from inpatient clinical information systems and provided a direct comparison between the two.

The study has several limitations. We assumed that all electronic medication records copied to the PAML were validated, and that the patients were not actually taking medications that were not copied to the PAML. It is possible, however, that PAMLs could also be inaccurate. If identical medication entries were present in more than one clinical information system, only one would be copied to the PAML, even if all were valid. However, the majority of medication entries analyzed in the study did not have duplicates, and in multivariable analysis of medication

record characteristics predictive of validation all of our findings were independent of duplicate records. Our study was limited to two academic medical centers in Boston; it is therefore possible that the findings may not apply elsewhere. We only studied patients who were admitted to the hospital, and the characteristics of the electronic medication records of the patients who were not admitted could be different. Finally, as with any retrospective analysis, the data were not collected specifically for the study which could potentially introduce bias in our findings.

In summary, this large retrospective study of validation of medication records in clinical information systems has confirmed and quantified findings previously reported in smaller investigations and demonstrated new relationships between the characteristics of electronic medication records and their validity. These findings could be incorporated in the design of the future clinical information systems to ensure that information predictive of non-validity of medication records is prominently displayed and that special procedures are devised for handling classes of medications (for example, antibiotics or *prn* medications) that are particularly likely to be outdated.

Acknowledgments

We would like to thank Partners Medication Reconciliation team for their hard work that made this project possible. We would like to express our gratitude to James Shalaby and Jaime Chang for assistance with medication class mapping.

References

[1] Bates DW, Gawande AA: Improving safety with information technology. *N Engl J Med* 2003, 348(25):2526-2534.

[2] Manley HJ, Drayer DK, McClaran M, Bender W, Muther RS: Drug record discrepancies in an outpatient electronic medical record: frequency, type, and potential impact on patient care at a hemodialysis center. *Pharmacotherapy* 2003, 23(2):231-239.

[3] Ernst ME, Brown GL, Klepser TB, Kelly MW: Medication discrepancies in an outpatient electronic medical record. *Am J Health Syst Pharm* 2001, 58(21):2072-2075.

[4] Grant RW, Devita NG, Singer DE, Meigs JB: Improving adherence and reducing medication discrepancies in patients with diabetes. *Ann Pharmacother* 2003, 37 (7-8):962-969.

[5] Vilke GM, Marino A, Iskander J, Chan TC: Emergency department patient knowledge of medications. *J Emerg Med* 2000, 19(4):327-330.

[6] Vermeire E, Hearnshaw H, Van Royen P, Denekens J: Patient adherence to treatment: three decades of research. A comprehensive review. *J Clin Pharm Ther* 2001, 26(5): 331-342.

[7] Joint Commission 2006 National Patient Safety Goals Implementation Expectations http://www.jointcommission.org/NR/rdonlyres/DDE15942-8A19-4674-9F3B-C6AE2477072A/0/06_NPSG_IE.pdf

[8] Poon EG, Blumenfeld B, Hamann C, Turchin A, Graydon-Baker E, Baker G, McCarthy PC, Poikonen J, Mar P, Schnipper JL *et al*: Design and Implementation of an Application and Associated Services to Support Interdisciplinary Medication Reconciliation Efforts at an Integrated Healthcare Delivery Network. *J Am Med Inform Assoc* 2006, in press.

[9] Wagner MM, Hogan WR: The accuracy of medication data in an outpatient electronic medical record. *J Am Med Inform Assoc* 1996, 3(3):234-244.

[10] Curtis JR, Westfall AO, Allison J, Freeman A, Kovac SH, Saag KG: Agreement and validity of pharmacy data versus self-report for use of osteoporosis medications among chronic glucocorticoid users. *Pharmacoepidemiol Drug Saf* 2006, 15(10):710-718.

[11] Yang JC, Tomlinson G, Naglie G: Medication lists for elderly patients: clinic-derived versus in-home inspection and interview. *J Gen Intern Med* 2001, 16(2):112-115.

[12] Schnipper JL, Kirwin JL, Cotugno MC, Wahlstrom SA, Brown BA, Tarvin E, Kachalia A, Horng M, Roy CL, McKean SC *et al*: Role of pharmacist counseling in preventing adverse drug events after hospitalization. *Arch Intern Med* 2006, 166(5):565-571.

Address for correspondence

Alexander Turchin, MD, MS
Clinical Informatics Research and Development
Partners HealthCare System
93 Worcester Road, Suite 201
Wellesley, MA 02481
aturchin@partners.org

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Evaluation of an Electronic Medication Reconciliation System in Inpatient Setting in an Acute Care Hospital

Abha Agrawal[a], MD, FACP, Winfred Wu[b], MD, MPH, Israel Khachewatsky[a]

[a] *Kings County Hospital, Brooklyn, NY, USA*
[b] *Suffolk County Department of Health Services, Hauppauge, NY, USA*

Abstract

Background: Medication reconciliation (MedRecon) is being implemented in many healthcare facilities as a means to reduce medication errors. However, there is scant literature on the evaluation of electronic MedRecon systems.

Objective: To evaluate the rate and type of discrepancies between a patient's home medication history and admission orders and to analyze factors affecting their occurrence using an electronic MedRecon system.

Design / Methods: We analyzed 3,426 consecutive inpatient admission MedRecon events from August to October 2006 in an acute care hospital using a recently implemented electronic MedRecon system.

Results: Overall, discrepancy rate was 3.12% (n=107) with omission of a home medication being the most common type (56.52%, n=65) of discrepancy. Admission time (8 PM to 8 AM), and total home medications >4 were found to have a significant positive correlation with discrepancy rate.

Conclusion: Using multidisciplinary MedRecon process based on an electronic system, we found a low discrepancy rate between patient's home medication history and admission orders compared with the rate in the literature, implying that an electronic MedRecon system is an important tool for improving patient safety.

Keywords:

medication reconciliation, medication errors, adverse drug events, patient safety, JCAHO, computerized medical record, quality of health care

Introduction

Occurrence of medication errors in acute inpatient care facilities in United States is a widely published issue and prevention of medication errors has become a critically important national priority (1). Many of the medication errors occur during care transition points such as hospital admission, transfer and discharge due to multiple changes in medication regimens (2).

In response to these concerns, the Joint Commission on Accreditation of Healthcare Organizations (JCAHO), in 2006, mandated that all JCAHO-accredited facilities must "accurately and completely reconcile medications across the continuum of care."

The Medication Reconciliation (MedRecon) process on admission involves compiling a complete and accurate list of a patient's home medications and comparing that list to a provider's admission orders. Many health care facilities have implemented paper processes to meet the challenging task of performing effective MedRecon (3-5). Recent publications have reported some experience with electronic MedRecon systems (6-7); however the use of such systems is a relatively new phenomenon and evaluation of their effectiveness in preventing potential medication errors is of much importance as many hospitals are expected to implement such systems in the coming years.

The objective of this study is to evaluate the rate and type of discrepancies between a patient's home medication history and inpatient admission orders and to analyze the factors affecting their occurrence from the recently implemented electronic MedRecon system at our facility. Published research strongly suggests that a low rate of such discrepancies would imply a low risk of potential adverse drug events and a safer hospitalization (8).

Materials and methods

Setting and participants

The study was conducted at Kings County Hospital Center (KCHC), a 630-bed acute tertiary care teaching hospital with approximately 24,000 discharges per year. The inpatient services included in the study were medical, surgical, behavioral health, obstetrics / gynecology and other specialty services. Table 1 describes the patient, clinician, and environment of care characteristics for the 3,426 MedRecon events in the study.

Electronic Medication Reconciliation System at KCHC

In 2006, an electronic MedRecon system was implemented facility-wide as a module within the existing commercial electronic medical record (EMR) system, MISYS. The EMR is

Selected for best paper award.

Table 1 - Patient, clincian, and environment of care characteristics

Total n=3,426 unique MedRecon events	
Patient Characteristics	
Age, mean± SD, y	38.4 yr +/- 23.2
Sex	
Male	1418 (41%)
Female	2008 (59%)
Insurance status	
Insured	2686 (78%)
Uninsured	740 (22%)
Clinician Characteristics	
Resident	2405 (70%)
Attending physician	804 (23%)
Physician Assistant/Nurse Practitioner	217 (7%)
Environment of care characteristics	
Admission day	
Weekday	2631 (77%)
Weekend (Saturday or Sunday)	795 (23%)
Admission time	
Daytime, 8 AM or later	1584 (46%)
Night time, 8 PM or later	1842 (54%)
Service type	
Non-pediatrics	589 (17%)
Pediatrics	2837 (83%)

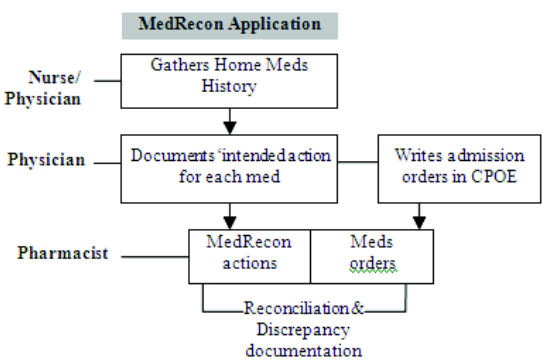


Figure 1 - Schemata of Admission Medication Reconciliation

extensively used across KCHC with all outpatients and inpatients orders being placed using the computerized physician order entry (CPOE) application in the EMR.

To comply with the mandatory JCAHO requirement of MedRecon, a dedicated multidisciplinary team consisting of physicians, nurses, pharmacists, information technology personnel, and clinical educators was convened. The team was responsible for the design, development, implementation, and evaluation of the electronic MedRecon module. The outpatient MedRecon module was implemented in all 100+ clinics from December 2005 to February 2006. The inpatient MedRecon system

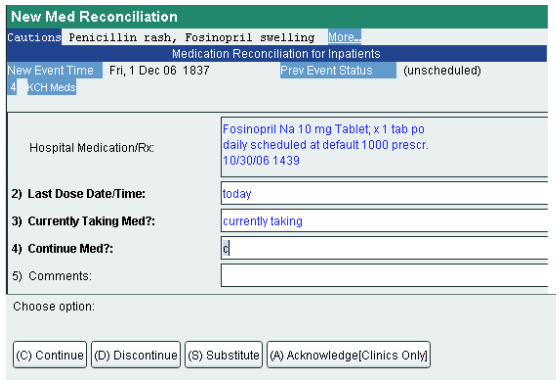


Figure 2 – Screen Shot Displaying Physician Medication Reconciliation Documentation Screen

was implemented for all services in May 2006. The inpatient system incorporates MedRecon processes for all three stages: admission, transfer and discharge. The current study evaluated the admission MedRecon process only, described below, as a majority of medication errors have been found to occur on admission (9).

The admission MedRecon process involves three steps (Figure 1). The first step consists of documenting a comprehensive medication history of a patient from all sources and can be performed by a nurse or a physician. The clinician invokes the MedRecon application on every admission by clicking on a MedRecon menu button in the EMR. The system automatically compiles a list of medications preexisting in that patient's profile (including outpatient or discharge prescriptions written using CPOE) in the EMR and presents them on the screen. These medications are labeled as "KCHC meds". For each KCHC med, the clinician documents the "currently taking" status (yes or no), last dose date and time, and optionally, any relevant comments. To ensure a complete home medication history, the next screen asks the clinician to add any medications being prescribed outside KCHC and non-prescription substances, such as Over-the-counter medications, nutraceuticals, herbals etc., labeled "Non-KCHC meds". The clinician documents "currently taking" status, last dose date and time, and an optional comment for "Non-KCHC meds" as well.

In the second step, a physician documents the 'intended action' for each medication in the MedRecon application by selecting one of the options: "continue", "discontinue", "substitute", or "unable to verify" (Figure 2). The MedRecon documentation is then automatically routed to an electronic work queue for pharmacy. The physician also places admission medication orders using CPOE. At the time of order writing, patient's home medication history, derived from the MedRecon application is made available in real-time to the provider, on the CPOE screen, to minimize discrepancies between home medication history and admission orders.

In the final step, a pharmacist performs 'reconciliation' by comparing the physician's 'intended action' for home medications with the admission orders. The pharmacist screen is configured to display both MedRecon documentation and CPOE admission orders side-by-side using the split-screen technique. The pharmacist records whether any discrepancies were found (yes or no) and if yes, categorizes them into one or more of the following: "continued" medication(s) not ordered, "discontinued" medication(s) ordered, dose discrepancy, frequency discrepancy, therapeutic duplication, and other (show screen shot). The taxonomy of discrepancies was derived from the literature (10-11). In addition, the pharmacist also communicates with the provider to resolve any discrepancies found.

Study design

The study was conducted as an observational study based on retrospective data analysis, performed as a quality improvement project. For data collection, an automated query was performed in the EMR to retrieve medication reconciliation data on all inpatient admissions from all services from August 1 2006 to October 30 2006. The data were analyzed for overall discrepancy rate, type of discrepancies, and the association of various patient, provider and environment of care characteristics with discrepancy rates.

Statistical analysis

Statistical analysis was performed using Microsoft Excel and SPSS for Windows, version 13.0. Chi-square analysis was used for comparisons of proportions. Patient, clinician, and environment of care characteristics were also analyzed using logistic regression. These characteristics were entered into the model and adjusted odds ratios with 95% confidence intervals were calculated. All P-values were 2-sided and a significance level of 0.05 was used.

Results

A total of 3,426 consecutive unique MedRecon events were found in the system during the study period. Of these, 107 events had at least one discrepancy recorded, yielding an overall rate of 3.12%. The categorization of discrepancies is described in Table 2. A total of 115 types were recorded as one event may have more than one type of discrepancy recorded. Unintentional omission of a home medication from admission orders was the most frequent type of discrepancy.

Table 3 summarizes various characteristics according to the presence of discrepancy based on bivariate analysis. Table 4 describes the association between selected variables and discrepancies based on multivariate analysis. Total medications >4 and night time admission were significant positive predictors of discrepancy and weekend admission was a significant negative predictor.

Table 2 - Categorization of discrepancies found

Type of discrepancy	No.	%
"Continued" medication not ordered	65	56.52%
"Discontinued" medication ordered	12	10.43%
Dose discrepancy	11	9.57%
Frequency discrepancy	1	.087%
Therapeutic discrepancy	4	3.48%
Other	22	19.13%
Total	**115**	

Table 3 – Patient, clinician and environmental characteristics according to the presence of discrepancy

	Discrep. present	Discrep. absent	p value
Patient Characteristics			
Age >65	28%	14%	<0.01
Uninsured	20%	22%	0.61
Total meds >4	31%	13%	<0.01
Clinician Characteristics			
Resident	80%	70%	0.02
Nurse didn't perform meds history	59%	66%	0.11
Care Environment Characteristics			
Weekend admission	14%	24%	0.02
Nighttime admission	69%	53%	<0.01
Pediatrics service	13%	17%	0.25

Table 4 - Association between selected variables and discrepancies

	Adjusted OR*	95% CI*
Patient Characteristics		
Age >65	1.59	0.99-2.55
Uninsured	1.08	0.66-1.76
Total meds >4	2.25	1.41- 3.57
Clinician Characteristics		
Resident	1.57	0.94 - 2.62
Nurse didn't perform meds history	1.23	0.81-1.88
Care Environment Characteristics		
Weekend admission	0.50	.284-.863
Nighttime admission	1.84	1.20-2.81
Pediatrics service	1.03	0.55-1.95

* OR - Odds Ratio; CI- Confidence Interval

Discussion

Current study reports a relatively low rate (3.12%) of discrepancies between a patient's home medication history and admission orders. Other studies have reported that 54-67% of all admitted patients have at least one discrepancy

between home medications and the actual admission orders (8, 12-14).

Similar to reported literature, we found that error of omission is the most common prescribing error on admission, followed by error of commission. We also observed that about 20% of the discrepancies were recorded under the free text category of 'other'. The contents of this field are being analyzed to understand if another category of discrepancy should be included in the system.

A earlier preliminary 2-week (June 3-June 16, 2006) investigation of MedRecon events at our facility using the aforementioned system, shortly after implementation, had found a 20% rate of discrepancies (15). This significant reduction in the discrepancy rate over the next 4 months implies that the system is remarkably effective in preventing discrepancies as its implementation has matured. Reduction of discrepancies is an important factor in improving patient safety as a recent analysis reported that discrepancies have the potential to cause moderate to severe discomfort or clinical deterioration in 38.6% cases (8).

Our study is novel in reporting not only the discrepancy rate using an electronic MedRecon system but it also reports various factors influencing the discrepancy rate. On adjusted analysis, two factors were found to have a significantly positive correlation with the occurrence of a discrepancy – night time admission and total home medications >4. This finding underscores the risk of polypharmacy and reiterates the need to exercise extra caution in writing admission orders during night hours. Weekend admission was found to have a significant negative correlation with the occurrence of discrepancies. This may be secondary to reduced workload on weekends, but further investigation is required to ascertain the implications of this finding.

This study demonstrates that although CPOE decision support tools are an important component of medication error prevention strategies, they alone are not sufficient to prevent errors of prescribing. An effective MedRecon process should complement CPOE systems in preventing medication errors for the following reasons. First, similar to prior research, we found that the most common error is omitting a medication from admission orders that is taken at home (10-11). CPOE systems with decision support tools such as drug-drug interactions, dose-range checking, therapeutic duplication etc. have proven quite effective in addressing the issues of lack of prescriber's knowledge about drugs; however, a CPOE system would not be capable of detecting unintentional omission of medications taken before admission. Second, CPOE, unless linked to a community pharmacy database, only 'knows' the medications prescribed at the primary institution and is 'unaware' of medications prescribed elsewhere. An effective MedRecon process, on the other hand, ensures a comprehensive home medication history. Third, CPOE profile only indicates that a medication was prescribed and not whether the patient is actually taking the medication – another issue addressed effectively with MedRecon solution.

The discrepancy rate in our study essentially highlights the 'safety gap' between gathering the home medication history list and (via MedRecon) the admission ordering (via CPOE), even when both processes are computerized. To eliminate medication errors, this gap - though small as reflected in our low discrepancy rates – needs to be bridged. One potential solution to bridge this gap is to allow clinicians to convert entries in home medication history into orders easily 'with a click'. Clinicians at our institutions expressed an interest in this solution as it would save them time. We deferred linking MedRecon directly to CPOE, at least in the first phase, for several reasons. First, several reports on unintended consequences of information technology caution that over-automation of workflow may introduce new errors (16-17). Second, a standardized normalization and data mapping approach would be required before converting outside home meds into CPOE entries. Therefore, currently we decided to keep home medication history detached from the CPOE, although it is always available on the order entry screen as a reminder to the physician.

The current study also highlights the important role of clinical pharmacists in preventing potential adverse events at admission. A recently published systematic review reported that the addition of clinical pharmacists to care teams resulted in improved care (18). Most studies in this review focused on their role in medication history taking or as a liaison service. Our study is unique in comprehensively evaluating their active role in the actual reconciliation process.

A useful 'side effect' of the medication reconciliation implementation at our facility has been a heightened awareness among prescribing physicians of the importance of a comprehensive home medication history. In addition, the system provides a standardized method of documenting outside medications and sharing the information among various providers across the continuum of care.

A limitation of our study is that no chart review was done to ascertain the impact of the discrepancies on clinical outcomes or length of stay. The study would also be strengthened by categorizing each instance of discrepancy into a severity score according to the potential to cause harm (12).

Conclusion

Using a multidisciplinary MedRecon process based on an electronic system, we found a low discrepancy rate between patient's home medication history and admission orders. Further research is needed to evaluate the impact of these discrepancies on clinical outcome. As informatics infrastructure for data mapping among disparate systems advances, integrating MedRecon system with CPOE should be explored.

Acknowledgements

The authors wish to thank other key members of the MedRecon project team at Kings County Hospital – Munish Khaneja, MD, MPH, Eugene Edynak, MD, Christopher Russo, RPh.

References

[1] Aspden P, Institute of Medicine (U.S.). Committee on Identifying and Preventing Medication Errors. Preventing medication errors. Washington, DC: National Academies Press; 2006.

[2] Bates DW, Spell N, Cullen DJ, Burdick E, Laird N, Petersen LA, et al. The costs of adverse drug events in hospitalized patients. Adverse Drug Events Prevention Study Group. JAMA 1997;277(4):307-11.

[3] Bartick M, Baron D. Medication reconciliation at Cambridge Health Alliance: experiences of a 3-campus health system in Massachusetts. Am J Med Qual 2006;21(5):304-6.

[4] Haig K. Medication reconciliation. Am J Med Qual 2006;21(5):299-303.

[5] Pronovost P, Weast B, Schwarz M, Wyskiel RM, Prow D, Milanovich SN, et al. Medication reconciliation: a practical tool to reduce the risk of medication errors. J Crit Care 2003;18(4):201-5.

[6] Hamann C, Poon E, Smith S, Coley C, Graydon-Baker E, Gandhi T, et al. Designing an electronic medication reconciliation system. AMIA Annu Symp Proc 2005:976.

[7] Poon EG, Blumenfeld B, Hamann C, Turchin A, Graydon-Baker E, McCarthy PC, et al. Design and Implementation of an Application and Associated Services to Support Interdisciplinary Medication Reconciliation Efforts at an Integrated Healthcare Delivery Network. J Am Med Inform Assoc 2006;13(6):581-592.

[8] Cornish PL, Knowles SR, Marchesano R, Tam V, Shadowitz S, Juurlink DN, et al. Unintended medication discrepancies at the time of hospital admission. Arch Intern Med 2005;165(4):424-9.

[9] Tam VC, Knowles SR, Cornish PL, Fine N, Marchesano R, Etchells EE. Frequency, type and clinical importance of medication history errors at admission to hospital: a systematic review. Cmaj 2005;173(5):510-5.

[10] Lau HS, Florax C, Porsius AJ, De Boer A. The completeness of medication histories in hospital medical records of patients admitted to general internal medicine wards. Br J Clin Pharmacol 2000;49(6):597-603.

[11] Beers MH, Munekata M, Storrie M. The accuracy of medication histories in the hospital medical records of elderly persons. J Am Geriatr Soc 1990;38(11):1183-7.

[12] Gleason KM, Groszek JM, Sullivan C, Rooney D, Barnard C, Noskin GA. Reconciliation of discrepancies in medication histories and admission orders of newly hospitalized patients. Am J Health Syst Pharm 2004;61(16):1689-95.

[13] Akwagyriam I, Goodyer LI, Harding L, Khakoo S, Millington H. Drug history taking and the identification of drug related problems in an accident and emergency department. J Accid Emerg Med 1996;13(3):166-8.

[14] Bikowski RM, Ripsin CM, Lorraine VL. Physician-patient congruence regarding medication regimens. J Am Geriatr Soc 2001;49(10):1353-7.

[15] Wu W, Khaneja M, Agrawal A. Evaluation of Electronic Inpatient Medication Reconciliation Implementation at an Acute Care Hospital [Poster]. In: 17th Annual Symposium on Health Care Services in New York: Research and Practice; 2006; New York, NY; 2006.

[16] Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. J Am Med Inform Assoc 2004;11(2):104-12.

[17] Koppel R, Metlay JP, Cohen A, Abaluck B, Localio AR, Kimmel SE, et al. Role of computerized physician order entry systems in facilitating medication errors. Jama 2005;293(10):1197-203.

[18] Kaboli PJ, Hoth AB, McClimon BJ, Schnipper JL. Clinical pharmacists and inpatient medical care: a systematic review. Arch Intern Med 2006;166(9):955-64.

Address for correspondence

Abha Agrawal, MD, FACP
Kings County Hospital Center
451 Clarkson Avenue, Brooklyn, NY 11203
E-mail: abha.agrawal@nychhc.org

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Computerized Management of Chronic Anticoagulation: Three Years of Experience

Beatriz H. Rocha,[a] Laura H. Langford,[a] Steven Towner[b]

[a] *Medical Informatics, Intermountain Healthcare, United States*
[b] *Department of Internal Medicine, Intermountain Healthcare, United States*

Abstract

Chronically anticoagulated patients taking the drug Warfarin require time intensive management and follow-up processes to avoid complications. The "Chronic Anticoagulation Clinic" (CAC) protocol is a set of production rules that help manage, treat, and follow-up such patients. The CAC protocol has been in regular use at Intermountain Healthcare (Salt Lake City, UT, USA) for over three years. The results demonstrate an improvement on the number of patients with anticoagulation levels within the desired target range. The protocol alerts have a high acceptance rate (83.4%) and were able to help patients remember to collect their next coagulation test. The CAC protocol results show that production rules can improve the management of chronically anticoagulated patients. Additional studies are required to verify if this experience can be transferred to other institutions.

Keywords:

clinical decision support systems, computerized protocols, anticoagulation, Warfarin

Introduction

Chronically anticoagulated patients are those that receive oral anticoagulation treatment with Warfarin for a long period of time. Medical reasons for chronic anticoagulation therapy include diseases such as Stroke, Atrial Fibrillation, and Deep Vein Thrombosis (DVT). Many published studies highlight the importance of properly managing these patients in order to avoid complications of excessive or sub-therapeutic Warfarin dosage [1, 2]. Improper anticoagulation management results in medical complications such as embolism or bleeding, along with increased health care costs.

The proper dosage of Warfarin is monitored using the International Normalized Ratio (INR) coagulation test. Studies have shown that patients that are managed through specialized anticoagulation services have better control of INR and fewer complications [1, 2]. The objective is to maintain the patient's INR within a "target range" that is specific for the patient.

A specialized anticoagulation clinic was created at Intermountain in 2001. From its inception, this clinic has been supported by a paper protocol (set of production rules) that helps with the management, treatment, and follow-up of chronically anticoagulated patients. The intent of this paper is to report on the experience obtained by using a computerized version of this protocol for a period of over three years.

Materials and methods

Patients who are receiving Warfarin are referred by their primary care physicians to the Salt Lake Chronic Anticoagulation Clinic (CAC). CAC is a nurse practitioner-led telephone-based clinic responsible for educating patients about Warfarin usage, and also for managing their Warfarin dosage through periodic INR tests [3]. Patients come to the clinic just once a year for a scheduled educational session. INRs can be collected in any Intermountain laboratory, taking into account patient preferences and time availability. Providers are alerted electronically about test results which are then communicated to patients by phone.

The initial set of production rules was created in "paper format" by one of the authors (ST), taking into account his personal experience and pertinent literature. These rules were put in routine use at the clinic in 2001. The "paper format" rules were later refined through knowledge engineering sessions and implemented as computerized production rules by another author (BHR). The computerized rules are commonly known as the "CAC Protocol". Since its initial deployment, the CAC Protocol has had a few minor revisions to better integrate the rules with the workflow of the clinicians.

The CAC Protocol was implemented using "Foresight" [4] as its platform. Foresight is a flexible decision logic execution engine coupled with a sophisticated clinical data monitor, both developed at Intermountain Healthcare. Foresight is integrated with Intermountain's outpatient "Electronic Medical Record" (EMR) known as "HELP2" [5]. Patients referred to the clinic are enrolled in the CAC Protocol through HELP2. After enrollment, the clinician also enters in HELP2 the INR target range for the patient (e.g., "2.0 to 3.0" or "2.5 to 3.5").

Selected for best paper award.

Table 1 - Examples of CAC Protocol rules for the target INR ranges

Target INR	Current INR	Last INR (past 25 days)	Message	Severity	Method of Delivery
2.0 to 3.0	< 1.6	If no previous INR or previous INR was >= 1.8.	Action Point Low. Message if dosage of Warfarin was changed in the past 25 days. Inquire about signs and symptoms of clotting, and if necessary, refer to an appropriate facility for care. Consider extra dose of Warfarin. Increase weekly dose by 5-15%. Retest in 7-14 days. Previous INR zone: XXXXX (if available)	Medium	Message Log
	2.0-3.0	If previous INR was between 1.8 and 3.3 (and was a "Repeated Yellow Zone")	Possible sliding green zone. Message if dosage of Warfarin was changed in the past 25 days. Retest in 14-21 days. Previous INR zone: XXXX (if available)	Low	Message Log
2.5 to 3.5	5.0-8.9	If no previous INR or previous INR < 4.0	Action Point High. Message if dosage of Warfarin was changed in the past 25 days. Omit 1-2 doses. Inquire about bleeding and refer to appropriate facility for care if needed. MD to review. Retest in 24-48 hours. Previous INR zone: XXXXX (if available)	Critical	Message Log and Pager
	>= 9.0	N/A	Critical INR. Message if dosage of Warfarin was changed in the past 25 days. Hold Warfarin. Consider Vitamin K. Consult the medical director or nurse practitioner for advice. Retest in 24-48 hours. Previous INR zone: XXXXX (if available)	Critical	Message Log and Pager

The clinical data monitor activates Foresight every time a patient event (e.g., laboratory result, pharmacy order,

allergy documentation, etc.) is stored in the HELP2 database. The CAC Protocol rules are triggered every time a new INR result is stored into the HELP2 database. POC (point-of-care) INR

results can be entered directly into HELP2, activating the CAC Protocol in real-time.

Once triggered, the protocol rules start by verifying if the INR result is from an enrolled patient. If the patient is enrolled, the rules verify if the result is valid and if it has not been interpreted before, avoiding duplicated alerts. The rules also verify if an INR target range has been selected for the patient. If no target range can be found, an alert is generated requesting the clinician to select a target range.

Once these verifications have been completed, the rules interpret the new result by comparing it to both the last INR in the past 25 days and the last alert generated. Examples of the rules are shown in Table 1. During the interpretation phase, the new INR result is classified into a "zone" (Table 2). For instance, a "green zone" means that the new INR result is within the patient's selected target range while a "yellow zone" means that the new INR is slightly outside the target range (see Table 2 for possible zones).

Table 2 - Possible INR interpretation "zones"

Zones	Severity	Late alert
Critical	Critical	73 hours after 1st alert
Action Point High	Critical	97 hours after 1st alert
Action Point Low	High or Medium	21 days after 1st alert
Red (High or Low)	High or Medium	21 days after 1st alert
Possible Sliding Yellow (High or Low)	Medium	21 days after 1st alert
Repeated Yellow (High or Low)	Medium	28 days after 1st alert
First Yellow (High or Low)	Low or Medium	31 days after 1st alert
Possible Sliding Green (within target range)	Low or Medium	31 days after 1st alert
Green (within target range)	Low	50 days after 1st alert

At the end of the interpretation phase an alert message is generated classifying the current INR into a zone. The alert message also contains recommendations about dosage management, information if the Warfarin dosage has changed in the past 25 days, when the patient should collect the next INR, and the previous INR zone (if available). The alerts are also classified by severity (Table 1). A dif-

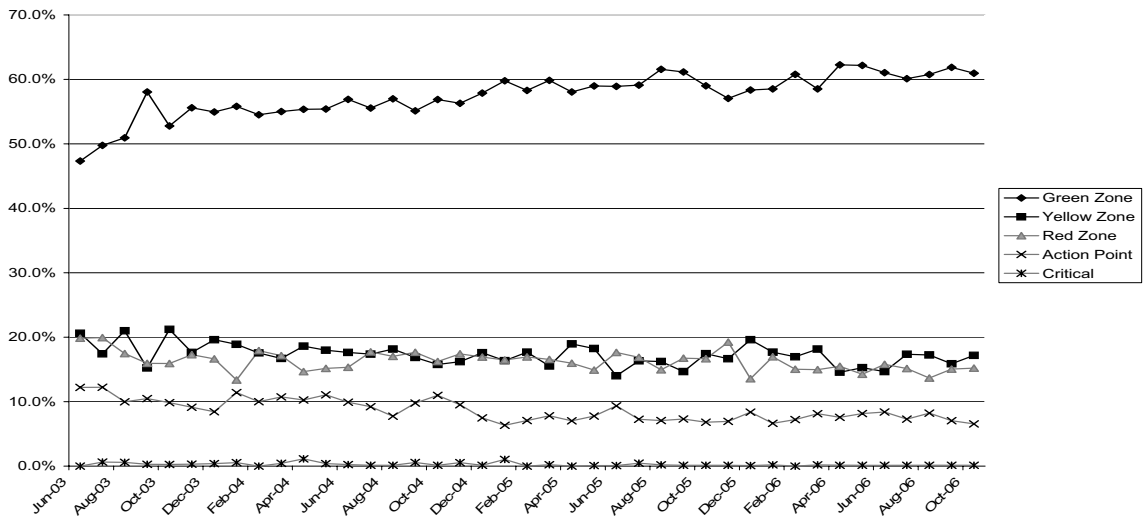


Figure 1 - Alert distribution by interpretation zone

ferent alert message is generated when an INR result is originated from an inpatient setting, since CAC clinicians do not manage Warfarin usage while patients are hospitalized.

The alert message is stored in the HELP2 database and also routed to the HELP2 "Message Log". Message Log is where the HELP2 users receive internal emails, notifications, and alerts. The alert message is sent to all clinicians that work in the CAC and that have an association with patient. Besides sending the alert to the Message Log, "Critical" and "Action High" alerts are sent to the pager of the clinician responsible for the clinic. If the alerts are not acknowledged within one hour, the alert message is resent to the same pager.

Clinicians use Message Log to review the CAC Protocol alerts that were issued since their last access. The effort to review the alert messages is divided among the clinicians, taking into account the severity of the alerts. The nurse practitioners (NPs) review "critical" and "high severity" alerts, while the licensed practical nurses (LPNs) review "medium" and "low severity" alerts. Besides the alert messages, the clinicians can also access from the same screen the INR result that triggered the alert. After reviewing the alerts, clinicians call each patient to communicate the INR result and, if necessary, the appropriate Warfarin dosage changes.

After contacting a patient, the clinician acknowledges the alert by accepting or rejecting it. If the clinician decides to accept the alert, she is presented with a set of possible actions taken, such as "Left Message with Household Contact" or "Spoke with Patient". The clinician can also enter other actions, but only as free-text. If the clinician decides to reject the alert, a set of rejection reasons is presented, such as "Perioperative management by another provider" or "Surgery/Procedure (Warfarin held)". Once an alert is acknowledged, it is automatically removed from the Message Log of all clinicians. If a clinician does not want to acknowledge an alert, she can add a comment to the alert such as "Called patient twice; not able to reach anyone at home". All alert acknowledgement details, including the clinician identity, the transaction time stamp, and the action taken are stored with the alert in the HELP2 database. Previously stored alerts can be reviewed at any time by the clinicians through the HELP2 Alert Review module.

During the INR interpretation phase, a "future event" is also created by Foresight depending on the resulting INR zone. A future event reactivates the CAC Protocol after a predetermined amount of time has elapsed. Each INR interpretation zone has a predefined "late" retest time period (Table 2). Whenever a new INR result becomes available before the expiration of the retest period, the future event is canceled and a new time event is created for the new INR result. If a new INR result is not stored before the predefined retest period expires, a "late alert" is generated reminding the clinician that the patient is due for a new INR test. Once the first late alert is generated, it is reissued once a week until a new INR result becomes available, or until the patient is un-enrolled from the CAC Protocol. Also, if the patient is enrolled in the CAC Protocol but an INR result is not made available, an alert is generated after 50 days reminding the clinician that the patient needs to have an INR test.

Results

The CAC Protocol was first deployed in June of 2003 and has been in production use ever since. The CAC Protocol has 46 rules total. The rules have been changed 11 times (9 times during the first 9 months of production use). One example of the changes made were the "late" retest time periods.

The Salt Lake CAC currently has 734 patients enrolled in the computerized protocol. A total of 48,552 alerts have

been generated between June of 2003 and October of 2006. An average of 52 alerts per day are generated during weekdays and 9 alerts per day during weekends. Figure 1 shows the distribution of alerts by interpretation zone. Alerts of the same type were aggregated to facilitate visualization, i.e., "green zone" includes "Green Zone" and "Possible Green Zone" alerts and "red zone" includes "Red Zone High" and "Red Zone Low" alerts.

Table 3 shows the number of alerts distributed by type. Of all patients ever enrolled in the protocol, 81.8% have received at least one "Patient late for follow-up INR" alert. The average number of late alerts is 8.1 alerts/patient (median is 5.0 late alerts/patient). All Salt Lake CAC alerts have been acknowledged: 40,492 (83.4%) have been "accepted" and 8,060 (16.6%) have been "rejected". The average time for acknowledging an alert after it has been generated is 13.6 hours (median is 3.0 hours). Only 3 patients, out of the 734 currently enrolled in the CAC Protocol, have not received an alert in the past 50 days. All 3 patients were recently enrolled in the protocol and are still in the initial 50 days waiting period.

Table 3 – Number of alerts by type

Type of alert	# of alerts	Percentage of total
Green Zone	15,966	32.9%
Patient late for follow-up INR	9,408	19.4%
Possible sliding Green Zone	4,328	8.9%
INR drawn for hospitalized protocol patient	3,896	8.0%
Red Zone High	3,369	6.9%
First Yellow Zone Low	2,968	6.1%
Action Point Low	2,475	5.1%
First Yellow Zone High	2,302	4.7%
Red Zone Low	2,293	4.7%
Action Point High	506	1.0%
Repeated Yellow Zone Low	447	0.9%
Repeated Yellow Zone High	252	0.5%
Patient INR goal has not been set. Please chart the goal.	200	0.4%
Critical INR	89	0.2%
Possible Sliding Yellow Zone High	27	0.1%
Possible Sliding Yellow Zone Low	26	0.1%
Total	48,552	100%

The most common reasons for accepting alerts were "Notified by phone" (23,975), "Spoke with Patient" (6,912), "Send this alert through voicemail to the patient" (4,386), and "Notified in person" (694). The most common reasons for rejecting an alert were "Hospitalization" (2,741), "Perioperative management by another provider" (578), "Deviation from the protocol" (437), "Noncompliance with Medication" (291), and "Incorrect Zone" (289). Out of the 289 alerts rejected as "Incorrect Zone" (0.6% of all alerts), 200 were "Patient late for a follow-up INR" alerts and 46 "INR drawn for hospitalized patient" alerts. The rest of the alerts rejected as "Incorrect Zone" were "Green Zone" (12), "Yellow Zone" (23), "Red Zone" (4), and "Action Point Zone" (4).

The Salt Lake CAC is run by two LPNs (one full-time and one part-time) and two NPs (both part-time), or 2.25 full-time equivalents (FTEs). Since 2005, 13 other CACs have been created within Intermountain. The 14 CACs are now responsible for the chronic anticoagulation management of 1,760 patients distributed within the state of Utah. All 14 clinics use HELP2 and the same computerized CAC Protocol.

Discussion

There are several publications describing the use of computerized decision support systems to help manage chronically anticoagulated patients [6-10]. These studies have shorter follow-up time and fewer patients when compared to the experience here reported. Most studies are related to Warfarin dose management [6,7,9,10]. However, all studies describe stand-alone systems that require users to enter pertinent patient data and also the recurring INR results. Some systems also suggest when the next INR should be collected [8-10]. No other study has reported the results of using a computerized oral anticoagulation decision support system fully integrated with an EMR, or where clinicians actively use the system to manage a large group of patients for a period of more than three years.

The results observed at the Salt Lake CAC confirm that a computerized protocol can help with the management, treatment, and follow-up of chronically anticoagulated patients. The CAC Protocol can be considered a relatively simple computerized protocol, given the small number of rules and the intent of handling only two INR target ranges. Despite the frequent changes to the rules during the first 9 months, which were primarily required for handling events that were not conceived during the initial design of the protocol, the CAC Protocol has been rather stable. Despite the simplicity of the protocol, its reliance on features of Foresight that are not commonly available in other EMRs makes its transferability to other healthcare institutions potentially difficult.

Figure 1 shows a steady improvement on the number of INRs maintained within the target range since the CAC Protocol was first implemented (from 47.3% to 60.9%). The proportion of patients with INRs within the target range is comparable or better than similar results observed in the literature [2,6,7]. It is conceivable that similar improvements could be obtained with a non-computerized version of the same protocol, but most likely not with the same staffing level given the relatively large number of enrolled patients. Furthermore, the clinic had been using the "paper format" protocol for two years and the observed improvements were after the implementation of the computerized version. The other 13 anticoagulation clinics that are now using the CAC Protocol will hopefully confirm the same improvements observed at the initial implementation site (Salt Lake CAC).

One of the main advantages of the CAC Protocol and its implementation using Foresight is that clinicians do not

need to know when and where the INR test was performed. Similarly, patients do not need to come for an in-person visit to the clinic and are able to go to the laboratory near their residence at their convenience. Clinicians also do not need to re-enter information that is already available in the EMR, including the recurring INR results. Once the INR result is stored in the EMR, clinicians can use the HELP2 Message Log to view the alert message and the triggering INR. Late INR collection alerts are also considered very useful, enabling proper management of all enrolled patients.

The alerts generated by the CAC Protocol had a high acceptance rate. The main reason for rejecting an alert was the fact that the patient had been hospitalized. These hospitalization events occurred outside the Intermountain network and were not recorded in the HELP2 database, i.e., the CAC Protocol interpreted the INR results as if the patient had not been admitted to a hospital. Since CAC clinicians do not manage the anticoagulation during hospitalization periods, they appropriately rejected the alerts. Conversely, the "Incorrect Zone" rejection reason had many different causes. The most common cause was lack of knowledge about the protocol rules by recently employed clinicians. The CAC Protocol was not designed to explain the reason why an interpretation zone was selected and new employees sometimes did not agree with the selection. These misinterpretations were normally solved with education, where a "paper" version of the CAC Protocol was made available to each clinician. The other two causes for "Incorrect Zone" alert rejections were patients that received late alerts but the clinician did not consider them late (e.g., INR test while patient was hospitalized, INR test done outside the Intermountain network and result not entered in HELP2), or INR results that were later corrected.

Due to the good experience with the current rules, a new set of rules was created for a new INR target range ("1.5 to 2.5"). These rules are being validated and should be implemented in January of 2007.

Acknowledgments

We would like to express our gratitude to Ilene Tippets, RN, MS for championing the implementation of the computerized protocol at the various clinics, and to Noman Rahman, David Steiner, and Randy Secrist for programming and maintaining Foresight.

References

[1] Witt SM, Sadler MA, Shanahan RL, Mazzoli G, and Tillman DJ. Effect of a Centralized Clinical Pharmacy Anticoagulation Service on the Outcomes of Anticoagulation Therapy. Chest 2005: 127 (5): 1515-22.

[2] Chiquette E, Amato MG, and Bussey HI. Comparison of an Anticoagulation Clinic With Usual Medical Care: Anticoagulation Control, Patient Outcomes, and Health Care Costs. Arch Intern Med 1998: 158: 1641-1647.

[3] Waterman AD, Milligan PE, Banet GA, Gatchel SK, and Gage BF. Establishing and running an effective telephone-based anticoagulation service. J Vasc Nurs 2001: 19(4): 126-32.

[4] Rocha RA, Bradshaw RL, Hulse NC, and Rocha BHSC. The Clinical Knowledge Management Infrastructure of Intermountain Healthcare. In: Greenes RA, editor. Clinical Decision Support: The Road Ahead. Boston: Academic Press, 2006.

[5] Clayton PD, Narus SP, Bowes WA 3rd, Madsen TS, Wilcox AB, Orsmond G, Rocha B, Thornton SN, Jones S, Jacobsen CA, Udall MR, Rhodes ML, Wallace BE, Cannon W, Gardner J, Huff SM, and Leckman L. Physician use of electronic medical records: issues and successes with direct data entry and physician productivity. AMIA Annu Symp Proc. 2005: 141-5

[6] Wurster M and Doran T. Anticoagulation Management: A New Approach. Dis Manage 2006: 9(4): 201-9.

[7] Fitzmaurice DA, Hobbs FDR, Murray ET, Holder RL, Allan TF, and Rose PE. Oral Anticoagulation Management in Primary Care With the Use of Computerized Decision Support and Near-Patient Testing. Arch Intern Med 2000: 160: 2343-48.

[8] Fihn SD, McDonell MB, Vermes D, Henikoff JG, Martin DC, Callahan CM, Kent DL, and White RH. A computerized intervention to improve timing of outpatient follow-up: a multicenter randomized trial in patients treated with warfarin. National Consortium of Anticoagulation Clinics. J Gen Intern Med. 1994: 9(3): 131-9.

[9] Margolis A, Flores F, Kierszenbaum M, Cavallo Z, Botti B, D'Ottone E, Tavella N, and Torres J. Warfarin 2.0--a computer program for warfarin management. Design and clinical use. Proc Annu Symp Comput Appl Med Care. 1994: 846-50.

[10] Ryan PJ, Gilbert M, and Rose PE. Computer control of anticoagulation dose for therapeutic management. BMJ 1989: 299: 1207-1209.

Address for correspondence

Beatriz H. Rocha, MD, PhD - Medical Informatics; Intermountain Healthcare; 4646 Lake Park Blvd; Salt Lake City, UT 84120; USA; email: Beatriz.Rocha@intermountainmail.org

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Physicians' Response to Guided Geriatric Dosing: Initial Results from a Randomized Trial

Josh F. Peterson, Benjamin P. Rosenbaum, Lemuel R. Waitman, Ralf Habermann, James Powers, Debbie Harrell, Randolph A. Miller[a]

[a]*Vanderbilt University School of Medicine, Department of Biomedical Informatics, Nashville, USA*

Abstract

Guided dosing within a computerized provider order entry (CPOE) system is an effective method of individualizing therapy for patients. Physicians' responses to guided dosing decision support have not been extensively studied. As part of a randomized trial evaluating efficacy of dosing advice on reducing falls in the elderly, CPOE prompts to physicians for 88 drugs included tailored messages and guided dose lists with recommended initial doses and frequencies. The study captured all prescribing activity electronically. The primary outcome was the ratio between prescribed dose and recommended dose. Over 9 months, 778 providers entered 9111 study-related medication orders on 2981 patients. Physicians using guided orders chose recommended doses more often than controls (28.6% vs. 24.1%, p<0.001). Selected doses were significantly lower in the intervention group (median ratio of actual to recommended 2.5, interquartile range [1.0,4.0]) than the control group (median 3.0 interquartile range [1.5,5.0], p<0.001). While physicians selected the recommended dose less than a third of the time, guided geriatric dosing modestly improved compliance with guidelines.

Keywords:

computerized physician order entry system; decision support systems, clinical; geriatrics.

Introduction

Dosing advice in computerized provider order entry (CPOE) systems frequently helps to individualize prescriptions [1-3] Systems have been developed to provide guided drug therapy based on age [2,4], renal function [5], and microbiological target [6]. Evaluations often show modest response rates to such offered recommendations. Iterative refinement of decision support systems may improve acceptance [7]. Yet, the evidence about what choices physicians will make when interacting with a guided dosing application remains empirical and less than definitive.

Previous work by author JP showed CPOE-based geriatric dosing guidelines can successfully improve inpatient fall rates [2]. The present study performed a randomized trial of guided dosing for patients 65 and older, to determine physicians' response rates.

Methods

Design of intervention: During care of elderly patients, the guided dosing system delivered advice to physicians about appropriate initial dosing for sedatives, neuroleptics, anti-emetics and skeletal muscle relexants for the most common indications. The geriatric dosing advisor also discouraged prescription of contraindicated drugs as defined by Fick and Beers (i.e. the "Beers Criteria") [8]. Because appropriate sedative and neuroleptic dosing ranges are broad for high-acuity (e.g., intensive care unit) patients, no barriers prevented selecting higher doses than recommended. The system utilized the same computational infrastructure that project members had previously implemented for specialized age- and weight-based pediatric dosing [4]. Figure 1 shows the editor screen that enables customization of dose lists, setting minimum and maximum single doses, and designating default doses for each study medication. Authors derived study-related dosing information for FDA-approved indications from "package insert" monographs, and for "off-label" indications from the medical literature and published textbooks. The CPOE-based text messages displayed along with study dosing information communicated titration strategies, possible adverse effects, and key monitoring parameters (Figure 2). An advisory group of 2 geriatricians [JP, RH], a geriatrics pharmacist [DH], and an internist [JFP] reviewed the doses and associated messages for accuracy. The knowledge base developed for this project will be submitted to a publicly accessible repository, the POGOe Geriatrics Web site sponsored by the D.W. Reynolds Foundation.

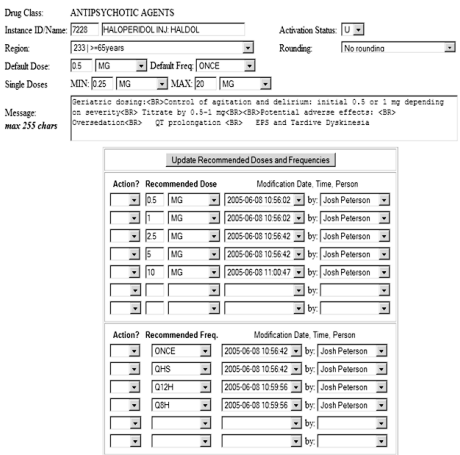


Figure 1 – Web editor for geriatric dosing knowledge base

Setting: The study took place in a tertiary care academic health center with 10 years experience with a self-developed CPOE system. Clinicians enter all inpatient medications orders into the CPOE system except for rare periods of downtime. At the study site, the majority of medication orders are entered directly by resident physicians (>70%), with the remaining 30% entered by attending MDs, nurses (verbal orders), pharmacists (transcribed written orders) and nurse practitioners.

Population: Patients 65 years and older receiving care on one of the order entry wards of the hospital including the emergency room, intensive care units, and a subacute unit were enrolled if admitted between December 8th and August 31st 2006. Patients who received no orders for an intervention were not analyzed. Likewise, only physicians caring for a study patient were analyzed. The CPOE system recorded all study orders including dosing parameters, ordering provider, and location of patient. Additionally, the system logged whether the physicians viewed the advice for intervention patients. For the analysis, study medication orders entered via order sets or specialized sedation protocols were excluded**.**

Analysis: The primary outcome was the ratio of prescribed dose over recommended dose. Medication orders with continuous frequencies (e.g., "q6h") were represented as a projected 24 hour dose and compared to the projected 24 hour dose of the recommended prescription. Single doses (e.g. "once only") were compared to the initial dose of the recommended prescription. Descriptive statistics were performed with medians and inter-quartile ranges (IQR); all dosing distributions were skewed. Significance tests for comparing intervention and control arms were performed with the Mann Whitney rank sum test. To assess for a crossover effect, where dosing advice on intervention patients influences decisions on control patients, we compared "control-only" physicians and "intervention-only" physicians and a pre-trial period to the trial period. All analyses were performed with R statistical package (http://www.r-project.org/).

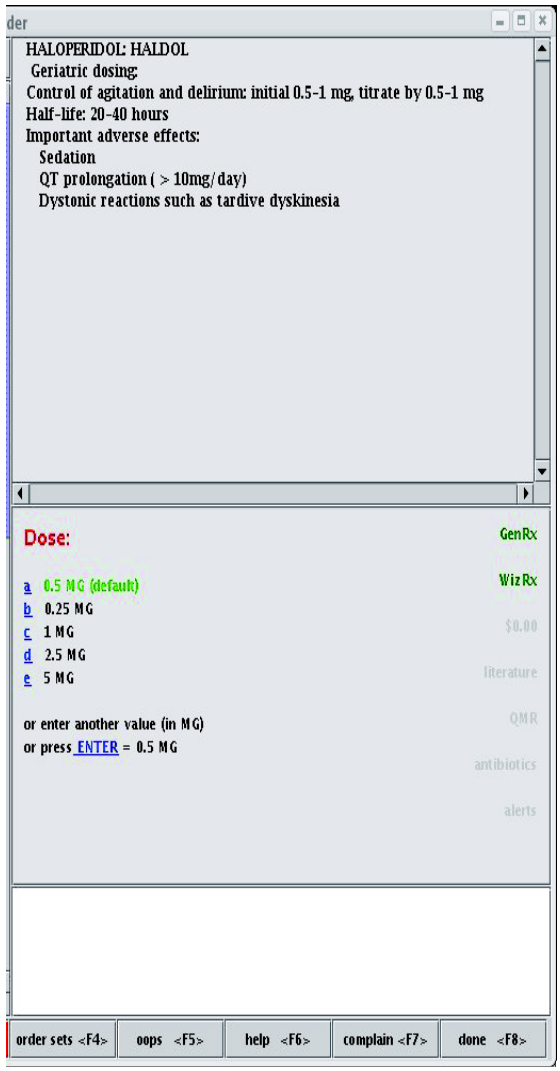


Figure 2 – Intervention integrated into institutional CPOE system

The institutional review board (IRB) approved the trial, including a waiver of consent for both patients and physicians.

Results

Over 9 months, 9111 study-related orders by 778 providers were entered for 2981 patients. Among the 88 study medications in the knowledge base, 23 were never ordered; many of these drugs were members of the Beers criteria list of potentially inappropriate medications and were not available from the hospital pharmacy. The overall acceptance rate of recommended doses was 28.6% in the intervention group vs. 24.1% in the control group, $p<0.001$). *Table 1* lists the median doses for all orders and predefined subgroups expressed as a ratio between the pre-

scribed dose and recommended dose. Overall, the intervention group of patients received lower doses than the control group (median 2.5, IQR [1.0,4.0] vs median 3.0, IQR [1.5,5.0], p<0.001). The difference comprised predominantly reduced doses of single-dose prescriptions, muscle relaxants, and slightly reduced benzodiazepine and anti-emetic prescription. The calculated 5th and 95th percentile dose ratios were 0.7 and 15 in the control group and 0.5 and 15 in the intervention group suggesting few differences at the extremes of the distribution.

In order to assess the effect of physician crossover, where physicians cared for both control and intervention patients, we compared the prescribed dosing from physicians who only cared for control patients (n=117) to physicians who only cared for intervention patients (n=103). Intervention-only physicians prescribed a significantly lower dose than control-only physicians (median 2.0 [1.0,4.0] vs median 4.0 [2.0,6.0], p<0001). The potential for crossover was also assessed by comparing a pre-trial period of 2 months (2315 medication orders) to the trial period. Pre-trial dosing was significantly higher with median 3.0 (2.0, 6.0) vs. median 3.0 (1.0, 5.0) for the trial period, p<0.001.

Beers criteria medications were prescribed in 34% of intervention prescriptions and 33% of control prescriptions. Dosing in the intervention group for medications in the Beers criteria list was not different than in the control group (Table 1).

*Table 1 – Median ratio of prescribed to recommended doses for control and intervention patients** IQR = inter-quartile range

Category	**Ratio of prescribed dose to recommended dose (median [IQR*])**		
	N	**Intervention**	**Control**
All Orders	9111	2.5 [1.0 , 4.0]	3.0 [1.5 , 5.0]
Drug Class			
Antihistamine/anti-emetic	2311	4.0 [2.0 , 4.0]	4.0 [2.0 , 6.0]
Benzodiazepines	2645	2.0 [1.0 , 4.0]	2.5 [1.2 , 4.2]
Neuroleptics	1473	4.0 [1.0 , 10]	4.0 [1.0 , 10]
Antihypertensives	1050	2.0 [1.0 , 4.0]	2.0 [1.0 , 4.0]
NSAIDs	442	4.0 [1.5 , 4.0]	4.0 [2.0 , 4.0]
Antispasmodics	292	2.0 [1.0 , 4.0]	3.0 [1.1 , 6.0]
Opiates	297	1.0 [0.5 , 1.5]	1.0 [0.4 , 1.5]
Sulfonylureas	305	4.0 [2.0 , 6.5]	4.0 [2.0 , 8.0]
Other anticholinergic	198	2.5 [2.0 , 5.0]	2.5 [1.0 , 5.0]
Other	98	1.0 [1.0 , 1.6]	1.3 [1.0 , 2.0]
Beers criteria medications	3051	2.0 [1.0 , 4.0]	2.0 [1.0 , 4.0]
Order Types			
Scheduled	5619	2.0 [1.0 , 4.0]	2.0 [1.0 , 4.0]
PRN	3492	4.0 [3.0 , 6.0]	4.0 [3.0 , 7.5]
Single dose	2580	1.0 [1.0 , 2.0]	1.25 [1.0 , 2.0]
Multiple dose	6531	4.0 [2.0 , 6.0]	4.0 [2.0 , 6.0]
Patient location			
Non-critical care unit	5028	2.5 [1.0 , 4.0]	3.0 [1.3 , 5.0]
Critical care unit and Procedure Suites	2463	3.0 [1.5 , 6.0]	3.0 [2.0 , 6.0]
Emergency Room	1338	2.0 [1.0 , 4.0]	2.0 [1.0 , 4.0]
Subacute unit	279	3.0 [1.5 , 6.0]	4.0 [2.0 , 4.0]

* IQR = inter-quartile range
Beers criteria medications are potentially inappropriate or with significant dose limitations as proposed by a consensus panel of geriatric experts

Discussion

This randomized trial demonstrated that physicians receiving advice from a guided geriatric dose advisor selected recommended doses for a minority of drug orders. Nevertheless, the guided dose advisor had a modest positive overall effect in decreasing the variation from recommended dosages for elderly patients. A low acceptance rate may be explained, in part, by the diverse indications for sedatives and neuroleptics in hospitalized patients. For example, benzodiazepines are used for both low acuity problems such as sleep as well as urgent indications such as combative behavior, pre-procedure sedation and acute alcohol withdrawal. Dosing may legitimately be higher for patients who have demonstrated tolerance or have chronically been taking a higher dose prior to admission. To improve physician acceptance and adherence to dosing recommendation, dosing guidance will have to cover a broader range of indications and design methods of capturing indication from physician users in order to route the user to a relevant dosing pathway. Single-doses were lower in the intervention group which may reflect a greater use of recommendations for "test-doses" which assess a geriatric patient's tolerance to a drug.

The differences between control and intervention groups were significant but of low magnitude. One explanation may be a "crossover" effect where physicians' learned of recommended dosing from treating intervention patients and made similar choices with control patients. Crossover was suspected because pre-trial dosing was higher than the control arm during the study, and "control- only" physicians prescribed doses significantly higher than "intervention-only" physicians. A crossover effect would bias the results towards the null. The trial was designed to randomize patients instead of providers because patient factors are most important when considering the primary patient outcome of the trial, falls. A crossover effect was considered unlikely during development because of the unobtrusive nature of the intervention. Potentially, the crossover effect represents evidence of physician learning. While this result will need more rigorous examination, if it is confirmed, clinical decision support systems may be serving an important teaching function in hospitals that implement CPOE.

The messages discouraging use of Beers criteria medications were generally ineffective with similar prescription rates in both control and intervention arms. Reasons for the lack of efficacy may include the lack of offered alternatives within the application or the inability for hospitalized patients to switch to preferred alternatives (e.g. due to allergies or intolerance). This outcome may also have been affected by the crossover effect mentioned above.

The interpretation of this trial was limited by several factors. Little information about the clinical context of the medication order was collected, and the overall appropriateness of the selected doses cannot be determined. There was no method to determine whether the physician making dosing decisions interacted directly with the guided dosing system. Some orders may have been dictated by proxy decision makers (such as senior residents or attendings) who were not directly exposed to the intervention. Finally, the influence of the specific CPOE system that hosted this project could not be determined. Compared to the CPOE system that hosted the previous implementation [2], the current system, has greater ease and fewer barriers to directly enter a higher dose which may encouraged physicians to bypass the dosing advice.

Conclusion

Guided medication decision support can affect geriatric dosing by modestly increasing agreement with guidelines. Higher levels of physician acceptance of recommendations will likely require greater use of indication-specific dosing advice.

Acknowledgments

Authors thank Osman Jalloh for his assistance with data collection. Dr. Peterson received support from the American Federation of Aging Research via the Merck/AFAR Junior Investigator Award in Geriatric Clinical Pharmacology. Drs. Miller, Peterson and Powers receive support from a Geriatrics Education grant from the D.W. Reynolds Foundation.

References

[1] Peterson JF, Bates DW. Automated selection of drugs and drug dose in patients with renal insufficiency. Medscape medication error and adverse drug reaction column 2002 March.

[2] Peterson JF, Kuperman GJ, Shek C, Patel M, Avorn J, Bates DW. Guided prescription of psychotropic medications for geriatric inpatients. Arch Intern Med 2005 April 11;165(7):802-7.

[3] Kuperman G, Anne Bobb, Bates DW. Medication-related clinical decision support in computerized provider order entry systems: a review. Journal of the American Geriatrics Society 2006.

[4] Waitman LR, Pearson D, Hargrove FR et al. Enhancing Computerized Provider Order Entry (CPOE) for neonatal intensive care. AMIA Annu Symp Proc 2003;1078.

[5] Chertow GM, Lee J, Kuperman GJ et al. Guided medication dosing for inpatients with renal insufficiency. JAMA 2001 December 12;286(22):2839-44.

[6] Evans RS, Pestotnik SL, Classen DC et al. A computer-assisted management program for antibiotics and other antiinfective agents. N Engl J Med 1998 January 22;338(4):232-8.

[7] Bates DW, Kuperman GJ, Wang S et al. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. Journal of the American Medical Informatics Association 2003 November;10(6):523-30.

[8] Fick DM, Cooper JW, Wade WE, Walker JL, Maclean JR, Beers MH. Updating the beers criteria for potentially inappropriate medication use in older adults. Archives of Internal Medicine 2003;163:2716-24.

Address for correspondence

Josh F Peterson, MD
Center for Health Services Research and
Department of Biomedical Informatics
Vanderbilt University School of Medicine
Ph: (615) 936-1465
josh.peterson@vanderbilt.edu

Chapter 7.

Usability

This page intentionally left blank

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Graphical Overview and Navigation of Electronic Health Records in a Prototyping Environment Using Google Earth and openEHR Archetypes

Erik Sundvall[a], Mikael Nyström[a], Mattias Forss[a], Rong Chen[a], Håkan Petersson[a], Hans Åhlfeldt[a]

[a] *Department of Biomedical Engineering, Linköpings universitet, Sweden*

Abstract

This paper describes selected earlier approaches to graphically relating events to each other and to time; some new combinations are also suggested. These are then combined into a unified prototyping environment for visualization and navigation of electronic health records. Google Earth (GE) is used for handling display and interaction of clinical information stored using openEHR data structures and 'archetypes'. The strength of the approach comes from GE's sophisticated handling of detail levels, from coarse overviews to fine-grained details that has been combined with linear, polar and region-based views of clinical events related to time. The system should be easy to learn since all the visualization styles can use the same navigation.

The structured and multifaceted approach to handling time that is possible with archetyped openEHR data lends itself well to visualizing and integration with openEHR components is provided in the environment.

Keywords:

information visualization; overviews; timelines; openEHR archetypes; medical informatics; medical records systems, computerized.

Introduction

Relating events in time plays an important role in getting an overview of a record. This paper first briefly provides an overview to and exemplifies some previous approaches to represent relationships between events and time. It then describes how Google Earth (GE) by simple means can be extended to become a prototyping environment capable of combining and extending the previous visualizations in a unified environment. New combinations and the use of time-lapse animation are also presented.

Currently some elementary script programming and XML authoring skills are needed to create the described visualizations, but we are extending the GE based environment with more tools and examples that will make it even easier to use by non-programmers. We have created Java based components that work 'behind the scenes' to provide easier translation from events in electronic health records (EHRs) to 2- or 3-dimensional space with optional time-based animation capabilities in the GE-based environment. Aggregation and summary functions are also handled by the components. The toolkit has a built in integration with openEHR[1] based data using the archetype approach to modeling, but could be connected to other data sources as well.

Towards end users this approach tries to exploit the ease of use in GE so that the user only needs to learn one way of navigating to use several different kinds of visualizations. Some functions are also provided to ease usability testing of developed visualizations. We are publishing our approach at an early stage, hoping to broaden the research community using it for future stages of further visualization development and usability testing. Color images and demonstrations illustrating the previous approaches and our prototyping environment can be found by looking up references, footnotes and the corresponding author's webpage.

Background

In the research field of Information Visualization, time series data have been of interest a long time. A number of approaches are exemplified below.

Linear time views

Improvements of overview and navigation of EHRs have been reported, e.g. by using graphical timelines, in the often cited LifeLines [1] project. The events visible on the timeline had short labels, and longer labels were revealed by pointing the mouse cursor at the event. The timeline view allowed visual correlation between events and also acted as a 'giant menu' that improved navigation since events could be 'clicked' to get direct access to detailed information in the EHR. Information categories, such as notes, medications, lab tests, were called *facets*[2] and were displayed as horizontal ribbons containing associated events. The user could control which facets should be open (showing events), or closed (showing a compressed 'silhouette' of the contained events without labels). Bade et al [2] explored visual timelines further and describe methods to show alerting qualitative levels in streams of quantitative data in compact ways. They also present methods for showing uncertainty of timepoint and value, trustability of data and periods of missing data. Further they suggest a method for interactive timeline distortion in order to decrease the screen space used by less interesting periods of time. Data from an intensive care unit was used to exemplify their approach and they included methods to compact high frequency data into efficient visual forms.

1 http://www.openehr.org/

2 The term facet will be used in the rest of the paper for information categories etc.

Polar time views

Time can also be represented in polar coordinates. One way is to do as in most 'analog' clocks where time runs clockwise around the circle. This has been explored in tulip plots[3], where different facets are represented as concentric arcs along the time circle. In order to investigate periodic patterns in time-series data, Carlis [3] and others have used spiral timelines where periodical data patterns surface visually if the period for a turn around the spiral corresponds to the periodicity in the data (e.g. yearly cycles of allergy symptoms).

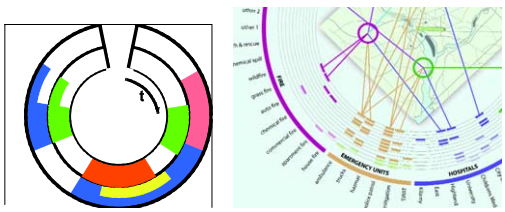


*Figure 1- **a** (left) A tulip plot[3] shows time clockwise and shows facets as concentric arcs. In Figure **b** (right) from [4] Livnat et al. put time radially outward and facets around the circle.*

Livnat et al. [4] use a polar coordinate system in an alternative way by putting the time axis radially outwards and distributing facets around the circle. Many polar representations leave space in the middle in order to avoid clutter. That space can be used to convey additional information. Livnat et al. put maps or network charts in the center and then relate the events around the circle to them using lines.

Time overlaid on maps and images

Kapler & Wright [5] describe a combined temporal and geospatial display, GeoTime, by putting a time (z-)axis perpendicularly upwards from a more or less flat (xy-)map. Events are shown further away from the map surface the further away they are in time. Kapler et al. also describe methods for aggregation of items based on proximity or defined regions.

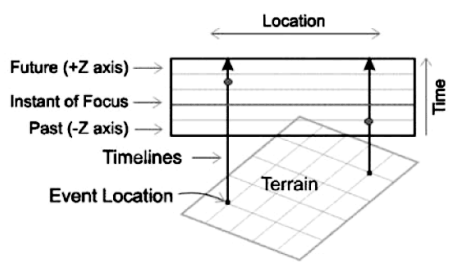


Figure 2 – GeoTime from [5]

Region based data entry

In Clinergy/Pen&Pad[4] [6] maps or charts of the body and organ systems are used as one way to specify parameters for data entry into an EHR. The user can point to coarse or detailed regions on the body charts and zoom in further to more detailed areas. The symptoms etc. available change depending on active region.

Visualization dimensions and variables

In [7] (section 4.3) parameters to use for visualization are summarized by Andrienko as being *dimensional* (referring to position in space and time plus various arrangements of display space) or *retinal* (size, shape, color, texture, orientation etc.) A recommendation is given that referrers of the dataset should be represented by dimensions and that attributes of the referrers should be represented by retinal variables.

Materials and methods

Archetypes

Archetypes are promoted by e.g. the openEHR foundation and the standards body CEN[5] as formalism to support modeling of clinical information structures and improved semantic interoperability between EHR systems. A detailed description of archetypes is outside the scope of this paper, see [8] for details. For the purpose of discussion they can be thought of as design descriptions of limited parts of the information structure in EHRs. Many such parts are then combined into hierarchies that constitute an EHR.

There are different aspects of time in a healthcare setting and for the purpose of recording they need to be treated carefully. The openEHR foundation has developed specifications for EHR systems and the topic of 'time' is extensively discussed in [9].

The information generated (e.g. using several *observation* archetypes) for a care event is collected into a *composition* that can be signed by clinical staff and put under version control. They can be flagged as being either *event compositions* that are intended for recording care events, or *persistent compositions* used for items of long term interest such as medication list, vaccination history, allergies etc. Compositions can contain entries of different kinds referring to events in the past (symptoms started a year ago) present or future (planned actions). In [10] *history* structures for *point events* and *interval events*, (possibly periodic) are defined. Methods for structuring e.g. long timeseries of measurements into compact summarized interval forms are described; this interval approach has some interesting similarities with summary approaches in [2].

3 Inspired by http://www.cas.lancs.ac.uk/alcd/visual/tulip_plots.html that refers to Barry, J.T. et al (1990) Graphical exploration of work history data. *Quad. Statist. Mat. Appl. Sci. Econ Sociali*., **12**, 65-74.

4 A Clinergy demo can be downloaded at http://www.opengalen.org/sources/software.html

5 EN13606, 'EHRcom', see http://www.centc251.org/

Google Earth

Google Earth[6] (GE) is an application that enables navigation of a digital globe with satellite imagery and map data. Developers can add new content in the form of XML files and images that can either be locally loaded into GE as files or published on a web server. If a server is used, then dynamic features such as reporting the current view of the user or updating previously loaded information are available.

Placemarks are icons with an optional text label that can be placed anywhere in a GE map. They can be clamped to the ground or put at an arbitrary height above the ground. When single clicking a placemark, a detailed description in the form of a 'text balloon' containing e.g. HTML-formatted text, possibly including images and hyperlinks, shows up. Double-clicking a placemark can change the 'camera' view of GE to a suitable position stored in the placemark by the placemark author. Hyperlinks in placemark description can be followed to related information that can either be a web page that opens in a built in or external web browser, or another set of GE objects (more images, placemarks etc.) that opens within the main GE navigation view. In this environment, these features can be used to go deeper and deeper into the EHR content. Transitions to the stored camera view in a placemark can be used as an efficient way to zoom in and position the view to a suitable angle that fits the next level of detail used in the EHR visualization.

Results

By placing an arbitrary image to be used as background over a piece of the globe we can hide the map and use that area as a 'desktop' for visualizations unrelated to the original map content. Initial tests show that most of the system features described in the introduction chapter can be reproduced in GE based visualizations. Linear and polar timelines can be added as images or drawn using built in GE geometric objects on top of the desktop.

We have also experimented with placement of notes that have relation to body parts (finding site of a tumor etc.) on organ system charts inspired by Clinergy. The notes are placed as placemarks above the relevant anatomical part in the chart and time is (as in [5]) represented by distance from the map. Just like real objects in a pile, old notes are at the bottom close to the map, and newer information is stacked above. By using the GE feature of region based loading and display of details, controlled by 'level of detail' (LOD) settings, different versions of a visualization piece can be shown depending on how close the user zooms in. If there are many notes in one region, they can be summarized into a single node indicating the number of notes when viewed at a distance, but shown as individual notes when zooming in. This has an effect similar to the region based aggregation discussed by Kapler et al. [5].

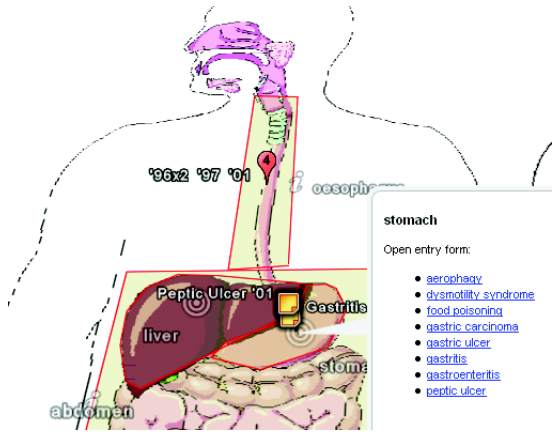


Figure 3 - A screenshot from Google Earth showing a tilted region based visualization in the prototyping environment. The abdomen region has been zoomed in so that subregions are visible. Two previous notes relating to stomach are seen, with the most recent one on top. The oesophagus region is too far away to show subregions and indicates that it contains four previous notes. A stomach related entry form selector has been opened by clicking the stomach 'target'.

Since it is possible to open web pages within GE, placemarks or information about current view can be used to select and open web based entry forms. This allows us to facilitate display and 'Clinergy-style input' in the same view. The usability of this combination remains to be tested. Maybe it is better to switch between dedicated viewing mode and a dedicated entry mode. Two-dimensional data like X-ray images or ECG plots are easy to show directly in GE. By using LOD, images with higher resolution can be fetched incrementally and shown when zooming closer to a low resolution thumbnail image.

If recent time is of most interest, then logarithmic time scaling can be used both for linear and polar diagrams. Related effects can also be achieved by simply tilting the view in GE when looking from recent time backwards. We have not yet attempted to develop ways to distort several different regions of a timeline simultaneously as in [2]. It would be interesting to compare this to the usability and understandability of regular GE-based navigation. Separate parallel timelines can be used for documentation events (versions of compositions) and clinical event timing (e.g. observations) if visual separation is desired. Important information (Patient ID, allergy warnings etc.) can be shown e.g. in a corner by using 'screen overlays' that are fixed to the screen and not to the map surface.

Facets and aggregation

How to section categories of information into facets like LifeLines [1] did with medications, lab tests etc. is an interesting research problem. In a (potentially) highly structured EHR e.g. based on archetypes we see these fairly simple computable divisions, based on:

6 http://earth.google.com/

- archetypes used, based on groups manually listed when creating the visualization, or by archetype class (observation, evaluation, instruction etc.) or external ontologies and systems [11] for classifying archetypes.
- openEHRs 'folder' structure that e.g. can be used for linking entries to health problems, care episodes etc.
- terminology system entities used in EHR entries e.g. ICD diagnosis codes or codes from SNOMED CT's main categories and/or grouping according to the target of relationships like 'finding site'.
- provider of the entry (organization, role or profession, person).

Aggregation that is needed e.g. in zoomed out views can be based on facets and/or time. In some cases (e.g. in the openEHR class 'History') some events may already have been manually summarized at the time of data entry and this could of course be used. Another obvious aggregation is over 'natural' time chunks like 'the number of entries per year'. In region based visualizations (body part maps etc.) geo-located aggregation based of the containment hierarchy of regions can be used. Better suited facets and aggregation can potentially be generated by decision rules and automated reasoning taking the above mentioned aspects and the current role and task of the user into account.

Time as a fourth dimension

In GE (from version 4) time is available as a fourth dimension. Objects (placemarks etc.) can be given time spans for validity. A time axis with sliders allows the user to select time span restrictions of what should be visible. By pressing the 'play button' the sliders can also be moved automatically which results in a time-lapse animation of changes over time. One way this can be used is to easily view previous states of the EHR in order to see how the medical picture grows over time. We want to explore if and how this can be used efficiently. We believe that a time-based view e.g. for the body/organ-system map where you can see information about problems etc. appearing (and possibly disappearing) over time may serve as a complement to using the z-axis to represent time. For medico-legal purposes it may also be useful to easily browse earlier states of the record. Clinicians will hopefully use the toolkit to find many other creative views and usages for time animations of EHR data – perhaps for studying the progress rate of diabetes complications etc. Users' changes of 'camera view' can be time-stamped and recorded. This information can be used e.g. in usability studies to record where most time has been spent etc. Time-lapse replay of frames showing the view trail in different ways is possible. Another possible use of recording can be to highlight for users which views in an EHR that have been most visited by other users—analogous to which pages of large paper based record that carry signs of frequent reading. Recording of view trails can also be used for medico-legal logging purposes.

Intended use

We do not expect that clinicians will use the toolkit in every day use; rather it can be used by them and others as a prototyping tool to invent and explore designs that can be used as parts of descriptions when ordering more polished and specialized systems from system providers.

Existing clinical images, either literally used (e.g. growth curves and partographs/partograms) or conceptually used (e.g. care-flows) can be used as a basis for fast prototyping of overviews provided that they can be captured as images.

We are aiming at decreasing the programming knowledge required to create visualizations. A user will be able to use GUI features to select the kind of visualization they want, e.g. linear, polar, region based (Clinergy inspired) and which axis to use for what (time, facets etc.), start and end points, width/radius etc. Connecting to openEHR data sources and selecting what data to fetch currently requires some scripting[7], but we are investigating the possibilities to select nodes using our Archetype Editor [12] or archetype visualization and browsing tools. A use case scenario could be as follows.

An overview based on a sketch of a tree of significant blood vessels and a linear timeline diagram for lab values is sketched on a whiteboard during a meeting. It is then captured by camera and the pictures are transferred into the GE based environment and anchored to the desktop on the map next to some previously developed overviews. Hierarchical regions containing each other are drawn and named on the 'vessel tree' and then mapped to entities in the archetype or terminology used in the EHR. For the lab result timeline the start- and endpoints on the time (x-)axis are marked in the image. Then the different lab value fields are mapped to positions for intervals and marked on the facet (y-)axis forming (possibly overlapping) ribbons along the diagram. Finally, color, icons, aggregation and summary strategies for the lab value plots are chosen.

Discussion

The technical solutions behind these extensions for GE are not very complex[8], instead it is more the possibilities for rapid development and evaluation that are of interest. Will the solutions be easy and efficient enough to be used by clinicians? If so, how will that affect the future development of interaction with EHR data? Will more efficient overviews and interaction possibilities be put into real clinical every-day use and would that have an impact on medical practice? Early feedback from clinicians and medical informaticians that have seen the current prototyping system has been positive and has resulted in comments like: "It opens up some new ways of thinking", "This would be nice when seeing new patients in primary care, but less useful for patents I already know well.", "[It] can change your cognitive level of interaction with the information" and "Information that is subtle or hidden beneath the surface might be found easier".

Future work

The prototyping environment presented in this paper is a prerequisite for planned user studies. Usability aspects of the visualization creation process will be one of the first

7 http://freemarker.sourceforge.net/

8 The approach can easily be reproduced by others. We intend to release a version of our solution as Open Source.

aspects to evaluate. Another aspect to study is the usability for end users of different created visualizations.

This paper focused on the 'dimensional' aspects of visualization – where to place information entities. How to create concise and efficient text labels has not been focused, neither have the 'retinal' parameters [7] (color, shape etc.). Systematic selection and use of icons and pictograms has not been the focus of our studies yet. Adding a structured approach to icon use along the lines of VCM by Lamy et al. [13] looks promising.

Our focus has been on viewing EHRs of individual patients one at a time, but the visualization principles would be applicable to studies of groups of patients as well.

The normal use of GE is to relate information to geographic positions, here we have not discussed this, but future work could include relating entry and retrieval of EHR content related to positions. In, e.g., ambulatory or distributed care EHR content and maps could be accessed in the same environment. We have started a location related experiment by using a map of a hospital ward as background image in GE and then placed the EHR visualizations for patients in their respective rooms.

If visualizations in GE are flat and the tilt function is not used, then the environment essentially becomes a zoomable 2D interface. Even though 3D visualizations often are more appealing Chen [14] (section 6.6) summarizes several 2D vs. 3D studies and concludes that increasing an interface from 2D to 3D is unlikely to improve the users task performance 'unless additional functions are provided so that users can have greater controls of objects in 3D interfaces'. Hence comparative studies of visualizations dependent on 2D and 3D respectively should be preformed. We believe that access to 'multi touch'[9] interfaces or hardware[10] dedicated to 3D interaction can affect user performance in 3D.

Conclusion

The capability and usability of geographical information systems of today like GE combined with the push for more structured and semantically well defined EHRs can in combination be used to create a powerful environment for prototyping overviews and interaction style for EHR systems.

We have summarized and unified approaches that may be used to create visualizations of temporal, casual and possibly anatomical relationship between events.

Acknowledgments

This work was performed in the framework of the EU-funded Network of Excellence entitled Semantic Interoperability and Data Mining in Medicine (SemanticMining). We want to thank SemanticMining network members, the OpenEHR community and Isabella Scandurra (Centre for eHealth, Uppsala University) for feedback during our work, Ocean Informatics for providing test data from their EHR Bank and the authors of [4] and [5] for permission to reuse illustrations.

9 E.g. http://cs.nyu.edu/~jhan/ftirtouch/

10 E.g. SpaceNavigator by http://www.3dconnexion.com/

References[11]

[1] Plaisant C, Mushlin R, Snyder A, Li J, Heller D, Shneiderman B. LifeLines: using visualization to enhance navigation and analysis of patient records. Proc AMIA Symp. 1998:76-80.

[2] Bade R, Schlechtweg S, Miksch S. Connecting time-oriented data and information to a coherent interactive visualization. CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems; 2004: ACM; 2004. p. 105-12.

[3] Carlis J, Konstan J. Interactive visualization of serial periodic data. UIST '98: Proceedings of the 11th annual ACM symposium on User interface software and technology; 1998: ACM; 1998. p. 29-38.

[4] Livnat Y, Agutter J, Moon S, Foresti S. Visual Correlation for Situational Awareness. INFOVIS '05: Proceedings of the 2005 IEEE Symposium on information Visualization; 2005: IEEE Computer Society; 2005. p. 13.

[5] Kapler T, Wright W. GeoTime information visualization. Information Visualization. 2005;4(2):136-46.

[6] Kirby J, Rector AL. The PEN&PAD data entry system: from prototype to practical system. Proc AMIA Annu Fall Symp 1996:709-13.

[7] Andrienko N, Andrienko G. Exploratory Analysis of Spatial and Temporal Data: A Systematic Approach: Springer 2005.

[8] Beale T. Archetypes: Constraint-based Domain Models for Futureproof Information Systems. OOPSLA 2002, workshop on behavioural semantics.

[9] Beale T, Heard S, Kalra D, Lloyd D. EHR Information Model; 2006.

[10] Beale T, Heard S, Kalra D, Lloyd D. Data Structures Information Model; 2006.

[11] Garde S, Knaup P, Hovenga EJS, Heard S (2007): Towards Semantic Interoperability for Electronic Health Records: Domain Knowledge Governance for openEHR Archetypes. Accepted for publication in Methods of Information in Medicine. http://healthinformatics.cqu.edu.au/downloads/Garde_openEHR_Methods_2007.pdf Accessed 2006-12-04.

[12] Sundvall E, Qamar R, Nyström M, Forss M, Petersson H, Åhlfeldt H, Rector A. Integration of Tools for Bind-ing Archetypes to SNOMED CT. In: Schulz S, Klein G, editors. Semantic Mining Conference on SNOMED CT; 2006 October 1 - 3; Copenhagen, Denmark; 2006.

[13] Lamy JB, Duclos C, Rialle V, Venot A. Which Graphical Approaches should be Used to Represent Medical Knowledge? Stud Health Technol Inform. 2005;116:719-24

[14] Chen C. Information Visualization: Springer 2004.

Address for correspondence

Erik Sundvall, Department of Biomedical Engineering, Linköping University, SE-581 85 Linköping, Sweden.
erisu@imt.liu.se,
http://www.imt.liu.se/~erisu/ (demo & images available)

11 Web links to references available at http://www.imt.liu.se/~erisu/

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Which Parts of a Clinical Process EPR Needs Special Configuration

Anders Barlach, Jesper Simonsen

Department of Communication, Business and Information Technologies, Roskilde University, Denmark

Abstract

Subject: *Which parts of an electronic patient record (EPR) can initially form a stable standard solution to be used by all clinicians? And which parts of an EPR can we predict needs initial as well as on-going re-configuration to meet the needs from diverse medical specialties.*

Purpose: *To analyze which screen types in a clinical process that can be standard configured and which are subject to initial as well as on-going re-configuration.*

Methods and results*: A pilot-project implementing a fully functional clinical process EPR was configured and used at a neurological ward, replacing all paper records 24/7. The analysis characterizes the different types of screens, a total of 243 included in the EPR solution. All screens have been extracted from the application and analyzed for changes in total 222 changes.*

Discussion and conclusion: *Most screens (87%) are very stable. Few (13%) are subjected to several re-configurations and they stabilize after an average of six iterations: Some may further stabilize over time since they address new but also general ways of working. Other screens relate to the specific medical specialty and cannot be part of a standard solution.*

Keywords:

integrated advanced information management systems, software design, interprofessional relations, user-computer interface, problem-oriented medical records

Introduction

Clinical Process form a core module of an Electronic Patient Record (EPR) that supports clinical documentation and decision making and comprises the on-going documentation of medical patient information made by the clinical staff (physicians, nurses, therapists, medical secretaries, etc.). It is hard to imagine that one single standard configuration would be optimal for all clinicians throughout the hospital: The clinicians at the neurological ward needs patient overviews focusing at parameters addressing cerebral haemorrhage and cerebral stroke while the clinicians from for example psychiatry needs completely different data and information in their patient overviews. Much basic patient data however (e.g. address information, family relations, drug profile, previous diagnosis, etc.) as well as functionality for common data entries (e.g. temperature, blood pressure, drug prescriptions, etc.) might be presented in a uniform way through standard screens used by all clinicians regardless of their medical specialties.

Modern EPR platforms today support international health standards (e.g. HL7), terminologies (e.g. SNOMED CT), and database platforms (e.g. Oracle HTB) while at the same time offering a high degree of configurability by means of e.g. XML-based templates. EPR technologies has reached a level where EPRs can be developed as a standard tool for all clinicians while the EPR at the same time can be configured to serve as a tool customized for specific needs supporting the high level of specialization that each clinical specialty represents. A main question that hospital managers and EPR developers face is thus the research question of this article: *Which Parts of a Clinical Process EPR Needs Special Configuration?*

The article is based on a project where a fully functional clinical process EPR was configured and used at a neurological ward, replacing all paper records 24 hours a day throughout a pilot lasting one week. We have analyzed how each of the EPRs 243 screens was configured and reconfigured to meet the needs of the clinicians. Based on this study we are able to indicate: 1) Which parts (screens) of an EPR that initially might be developed as a stable standard solution and used by all clinicians throughout many hospitals (no or few further re-configurations are to be expected when the EPR is rolled out to other wards); and 2) which parts of an EPR we can predict needs initial as well as on-going re-configuration to meet the needs from clinicians representing diverse medical specialties (e.g. neurology, surgery, gynaecology, psychiatry, etc.)?

In the following we introduce the study and the method used for the analysis. Then we present the results in terms of our categorization of screen types and the completed as well as requested re-configurations for each type. Finally we discuss and conclude our findings, the limitations of the study and the implications for practice and research.

The clinical process EPR project

The project was part of a research project on effects-driven IT development [1] (http://Effects-DrivenIT.dk) and was formed by 3 partners: Clinicians from the neurological stroke unit and project managers from the EPR unit at Roskilde County Hospital, researchers from the Department of Communication, Business and Information Technologies at Roskilde University and business architects from the vendor, CSC Scandihealth A/S. One main aim of the project was to experience how to configure a clinical process EPR module in participation with clinicians and to test how a configured solution would work in a real clinical process.

The project involved a neurological stroke unit treating patients with acute apoplexy where all paper-based patient

records were replaced with a configured and fully functional prototype EPR system. The aim was to evaluate an EPR with complete patient records tested on-line on real clinical processes [2, 3]. The project thus required thorough planning involving development of new EPR-supported patient trajectories, configuration and implementation of all screens needed in the EPR system, real-time integration with other systems, migration of patient data, and training of the clinical staff in using the system and working according to the revised patient trajectories.

The content of the EPR was identified during three workshops, i.e. the structure, content and placement of clinical notes and result templates, standard plans, concept lists etc. At the final workshop the complete specification was presented and reviewed before the actual configuration of the XML-based templates and load of the templates to the EPR. During this process the content of the EPR was elaborated in up to three iterative events. First, mock-ups were drawn on flip-over paper. Secondly, a preliminary non-interactive prototype was discussed. Finally, a running prototype was demonstrated, discussed, and evaluated. The vendor undertook the technical development of the prototype, along with interfaces to various legacy systems currently used at the hospital (ADT system, laboratory system, and medication module). A number of tests and re-configurations of the system were made in parallel with training the clinical staff in using the prototype. A final rehearsal was performed by testing the system under laboratory conditions using real patient-cases in a scenario setup on the solution that was due for release in the pilot. This was the final reassurance within the project team that the EPR was ready.

The most complicated part of the screens concerned those that should provide the clinical staff the ability to efficiently obtain overview and assessment of patients as well as on more efficient coordination in three specific and highly cooperative situations:

Nursing handover, which happens three times a day at the beginning of each nursing shift (7am, 3pm, and 11pm) and last about an hour. There is no time for the nurses that leave the ward to discuss patients with the nurses on the next shift. During the nursing handover, one nurse is designated as the team leader and provides an overview of the patients at the ward and manages the necessary coordination and exchange of information. This nurse reviews the patient records and orally informs the others about status and plans for the shift.

Team conference, which takes place once every weekday, lasts approximately 15 minutes, and includes all clinical staff members (physicians, nurses, and therapists). An inter-disciplinary assessment of each patient is carried out and plans are revised. The current status of each patient is given orally by a nurse and an overview of current plans is available by means of a table on a large whiteboard or, in the prototype EPR system, a full screen projected on the wall.

Medical ward round, which happens once every weekday and lasts for three to six hours. It includes evaluation, reviewing, and discharging of patients. The chief physician visits all patients and reviews the plans for their treatment. Usually there is no time for nurses to follow the physician during the ward round. Information exchange and coordination is obtained through the patient record and by ad hoc communication with the nurses on shift.

The required content was configured as XML-based templates that were loaded into the clinical framework tool, CSC Clinical Suite, based on the Oracle Healthcare Transaction Base (HTB). CSC Clinical Suite is not an EPR per se, but a clinical framework tool that can contain and present the clinical content as specified by the clinicians by use of XML-based tem-plates for overviews, clinical notes, results, standard plans, work situations and structure of the patients medical record. This makes it possible to configure a complete medical record in accordance with the clinicians requirements and is able to evolve dynamically as new requirements emerge.

In the final part of the project (the pilot), the configured EPR system was online 24 hours a day and replaced the paper-based records for all patients during one week in December 2005. Five years of patient data (in total more than 26 million data records from more than 300.000 patients) had been migrated to the EPR system and interfaces were established to the legacy systems in order to receive updated data during the project. The EPR system included screens projected on the wall during nursing handovers and team conferences, stationary and portable PCs, and PDAs used for obtaining measurements at the patients bedside (temperature, blood pressure, etc.). All clinicians used the EPR system during the pilot. Management oversaw to the project ensuring both legal requirements and patient-ethics were respected.

Data analysis method

All screens in the EPR solution 243 in total have been extracted from the application and analyzed for changes 222 in total made in the project period. The analysis is based on the vendors systematic documentation of all the changes made to each screen, from an initial first version of a screen and throughout the project period including the pilot where the system was used 24 hours a day.

In order to analyse the screens they have been categorized as follows. The screens have been divided into general and specific screens. A general screen can serve the same purpose on any medical ward: E.g. screens for recording basic vital values such as blood pressure, pulse and temperature might be the same on a medical and a surgical ward. Specific screens serve a special purpose within the given clinical speciality: E.g. screens for recording and monitoring a SIP score (Stroke In Progress) are specific to the Neurology speciality.

All screens in the EPR system (as well as in information systems in general) can be divided into two different categories, as either a form or a view:

- Form, resembling a paper form for *recording* (registration and submitting) data. This can be free-text or structured information in various degrees. Typical forms could be observations, notes, and basic vital values.
- View, is the *presentation* of data either recorded in the EPR system or received from external systems. A view retrieves data from one or several sources and presents it as information to the user or as an indexing service.

Typical information views could be: Graph presentation of basic vital signs, or overviews' creating information bulletin boards with focused information for a specific clinical situation (including e.g. nursing handover, team conference, and medical ward round). Views also include Journal structure views that present the user with all available data in a structure for navigation. This navigation hierarchy was designed to resemble that of the paper record.

In order to determine who made decisions regarding changes we assigned each view or form a primary user in terms of professional discipline (doctor, nurse, therapist or shared by doctor & nurse). This is to indicate the coordination involved among professional groups in the design and implementation process.

All changes analyzed were changes that were actually implemented. Several change request where also collected but not implemented because they were considered non-essential (nice-to-have as opposed to need-to-have) to the continued use of the system during the pilot. The changes made to the screens were analyzed with regard to when they occurred in the project (before, during or after completion of pilot). Types of changes include content (new fields in forms, new selections in views, labels changed); rules (business logic, validations); computations (adding or changing calculation functionality); and cancelled (retirement of screens due to time pressure or obsolescence due to other screens delivering similar services). The changes are summarized into 3 major groupings:

- None (0): No changes were necessary.
- Few and initial changes (1-2): One or two changes were made initially in the project during the prototyping process. These types of changes reflect a low complexity or uncertainties in design.
- Several and sustained changes (>2): More than two changes occurring including changes beyond the initial prototyping process. These changes reflect either uncertainty among clinicians or complexity in the implementations. It also reflect screens that needs to be configured by an *experimental* approach which entail several successive changes throughout the project, in some cases including changes made within the pilot period.

Based on the categories listed above all screens and changes were analysed and the resulting patterns are presented below.

Results

We have identified a number of interesting patterns with regard to the changes made to the screens representing the overall configuration (and re-configuration) of the clinical process EPR system. The implementation resulted in an EPR with a 4:1 ratio between forms and views. Less than 10% of the total 243 screens were specifically configured to the neurological specialty (16 out of 183 forms and 7 out of 60 views).

The majority of screens (87%) were not changed at all or only subject to few initial changes (table 1). Thus the major part of the total system may be considered as being quite stable. These stable screens were both medical specific forms (45%, 7 out of 16 totals) and general forms (90%, 152 out of 167 totals). Most of the stable forms were quite simple, in terms of containing only one or two data fields (e.g. registration of simple results like blood glucose) and often they were serving as a sub-template in larger and more complex forms. Views that present data from other systems were also very stable, e.g. views presenting X-ray results. Another characteristic for stable forms and views was that only one professional discipline was involved as main user, or the design was known from other systems as e.g. views presenting aggregated laboratory-results.

Distribution among screens changed

Total screens	None (0)	Few and initial (1-2)	Several and sustained (>2)
243	184	27	32
100%	76%	11%	13%

Table 1 - Changes made to the screens during the entire project.

The total number of changes accounts to 222. Out of these 83% (184 changes) were made to the 32 screens that received more than 2 changes each. This verifies that the configuration of 13% of the screens (32 out of 243) reflects a need for experimentation. These screens where subjected to a more thorough analysis and present interesting change patterns as seen below in table 2 and 3.

Screen change pattern specific vs. general

	Specific		General	
	Screens	Changes	Screens	Changes
Form	7	39	15	79
View	5	38	5	28
32	12		20	
184		77		107

Table 2 - Analysis of the 32 screens subjected to several and sustained changes (from table 1) distributed among screen requirements attributed to the specific neurology speciality or of a general clinical nature.

Screen change pattern among professional disciplines

	doctor	nurse	multi
Form	5	14	3
View	0	3	7
32	5	17	10

Table 3 - Analysis of the 32 screens to support either a professional discipline (doctor or nurse) or information collaboration among disciplines (multi).

Content changes where dominant in these patterns (82%, 184 out of 222 total changes), including adjusting labels on fields, adding new fields, removing obsolete fields, in some cases later to be added again. The need for experimentation grew according to the complexity of views, for example as more than one professional discipline was identified or data had to be drawn from several forms.

Motivations for changes in forms where often driven by their dependency to deliver data in the views. If for example a view is changed to include additional data this often entail that a form needs change in order to capture this data. Changing one view sometimes indirectly contributed to changes made in index-views that display a structure for navigating the various documentation models in the EPR. The scenario would typically be that each time a new view was available, it also had to be accessible without the search functionality, and this sometimes entailed that a logic entry or indexing in the Journal structure had to be assigned adding to changes accumulated by this index-view.

If we focus on the 32 screens from table 1 that were subject to several and sustained changes and display the results in table 2 and 3, we observe that the general forms are in majority to the general views (3:1). They are primarily owned by only one group of professionals (19 out of 22 have only one profession as primary owner). The nurses account for 11 out of 15 general forms shown in table 2 which also sparks attention to why they are represented with so relative many forms?

The views are equally distributed among specific and general, but are characterized by having more than one owner (7 out of 10, see table 3). Specific forms include forms for clinical plans with regard to stroke, which in the project constituted an entirely new way of applying their knowledge. These plans account for 4 out of 7 Specific Form screens shown in table 2, and 25 of 39 changes made to these screens.

Common factors contributing to changes among all the 32 screens has been identified as complex computations required on the client side, specific forms (e.g. Scandinavian Stroke Scale), or views (both specific and general) involving multiple professional disciplines where forms and views should support coordination of data or tasks. Especially when supporting an inter-disciplinary approach to EPR: E.g. complex views supporting the ward round or the team conference draw on information from radiology systems and clinical laboratory systems, and in addition including observations and notes made by doctors, nurses, and therapists.

The fact that the systems delivered 24/7 service during the pilot could entailed that only needed changes were implemented during the pilot (need-to-have changes as opposed to nice-to-have changes). During the pilot a few changes were deemed necessary in order to continue efficient operations. These changes occurred only to 3 views while the remaining 240 screens (99%) remained unchanged. Nevertheless the pilot and the use of the EPR in general were evaluated as being successful and measurements on clinical practice using the EPR has documented several significant improvements [1, 2, 3].

Discussion

The results indicate patterns of changes displaying themes that are predominant in the process of designing and implementing the clinical process EPR. Our study indicate that the majority of a clinical process EPR does not require special configuration with, regard to the different clinical specialties, as 87% of all screens in the EPR remained stable by requiring no or only few changes. The stable screens include simple forms, views presenting data from other known systems as well as forms and views addressing only one professional discipline. Content changes where dominant representing 82% of all changes. A substantial part of the changes is a result of chain reactions, typically where a change to a view subsequently trigger other changes in related forms or views for navigating the EPR. Screens with more than 2 content changes account for 143 out of 222 in total or 64% in only 29 screens. They were often related, e.g. FORM; Stroke Scale, Apoplexy Observations relate to VIEW; Apoplexy Overview AND FORM; Apoplexy Plan. Approximately 3 times as many forms were needed as views giving an idea of how many forms are required to sustain views.

A number of screens was subjected to several and sustained changes reflecting a need for an *experimental* approach to the process of configuring the EPR. The configuration of these parts of the system addresses application areas where the EPR introduce new ways of working. Potentially this might result in far-reaching improvements by ways of efficient support of inter-disciplinary coordination among multiple professional disciplines. We can predict that some parts of this configuration will stabilize over time since they address new but also *general* ways of working with EPR. Other parts of this configuration addresses themes related to the *specific* clinical specialty which indicate parts of EPR that can hardly be standard configured to serve clinicians throughout the hospital.

Forms supporting new ways of structuring documentation and views presenting the journal structure are examples of general parts of the EPR that faced several and sustained changes.

In our study the doctors applied their existing documentation model to the EPR and they retained dictating as usual with the medical secretary entering the dictate into the EPR. The nurses on the other hand, had to invent and specify their documentation model and integrate it with the doctors model in the journal structure (it was a deliberate part of the project to experiment with adding structure to the nurses documentation). This resulted in a higher activity regarding the design of forms with nurses as professional discipline (as indicated in table 3). Throughout the project this sparked several general discussions among the nurses about how they where using the paper based journal structure and how to use EPR. They could see a new perspective with the EPR, and the need to evolve their documentation models to include how they decode clinical data into nursing information. In general it was a challenged to figure out how to merge multiple documentation models serving their interdisciplinary needs without compromising their professional knowledge to accommodate other professionals. E.g. the doctors did not have to give up describing the paμtients anamnesis from the diag-

nostic perspective, just because the nurses would insist on describing the anamnesis from a holistic perspective.

Views presenting the journal structure support navigating the EPR and provide an alternative to the pre-defined information clustering implemented in the overviews. This also provides the users with the possibility to verify, in case of uncertainty, if they had missed some information in the overviews. They came to rely on the patient-journal structure for completeness. The upper levels in the journal structure must be general throughout the hospital and this we can be predict to become relatively stable over time (though this was not the case in our study where introducing clinical process EPR). However the lower parts patient-journal structure hierarchy might become more specialized and susceptible to changes thereby requiring occasionally experimentation and dynamic technological solutions for the medical specialties.

The parts of the clinical process EPR where an ongoing and experimental configuration can be identified as addressing the *specific* clinical specialty comprises support for highly cooperative activities such as planning the patient treatment and activities such as team conferences and nursing handovers.

Planning was the primary contributor to many changes in the doctors group. Planning account for 4 out of 7 specific forms having sustained changes and 25 of 39 changes made to the specific forms included among the top 32 screens listed in table 2. This was due to the fact that many of the planning and coordination tasks traditionally are handled by other professionals (nurses or secretaries). The story repeats itself as with the nurses lacking a documentation model, since the doctors had no prior system to rely on. Plans can overall be divided into 2 categories: The initiating or basic plan and the follow-up or supplementing plan. It was relatively easy to design the initiating plans as they to a high degree resemblance with the department guidelines. However the follow-up proved more difficult as they where often conditional (e.g. if X-ray result is positive order antibiotics) or involved coordination of tasks between professional disciplines or other medical specialities. This complexity is contributed to the innovation requirements of the professionals as they become aware of one-anothers areas of responsibilities and explore the possibilities of coordinating and sharing information in new ways.

The views supporting the coordinating activities during team conferences and nursing handovers were also subject to sustained changes. Although not many in numbers, they account to a significant number of changes: 33% of all changes listed in table 2 and 3. The changes were primarily content changes to views supporting interdisciplinary cooperation (team conference) or single disciplines being derived from an entirely new documentation model nursing observations.

Conclusion

The majority of screens (87%), were stable and include simple forms, views presenting data from other known systems as well as forms and views addressing only one professional discipline.

Relatively few screens (13%, or 32 out of 243) were subjected to several re-configurations and a part of these may further stabilize in the future since they address new but also general ways of working. Another part are screens specific to the clinical specialty. There are indications that only few specific screens are necessary per medical speciality.

The screens with sustained change requirements include both general and specific screens and comprise different types of views displaying the potential of an EPR: They present new ways of decoding and sharing information and supporting highly cooperative activities. These screens are characterized by the clinicians having no previous experience from a mainly paperbased everyday work environment, or clinicians involved in multi-disciplinary content and cooperative activity. Our project documents that such screens can be efficiently configured through an experimental and participative approach [4]. It is also clear that it requires continuing the experimental approach to include using the EPR in a real clinical everyday work environment. From the technological point of view it sets the standards for how the EPR vendors must be ready to meet the dynamic requirements and where to expect more confidence in the stability of the EPR.

The perspective of our study gives an indication as to what to expect when engaging in the implementation of a dynamic EPR. This paper present the result of just one pilot-test, and more tests are necessary to investigate the issues of the dynamic versus stable parts of a clinical process EPR. We are now applying our experience from the pilot to new projects where several medical specialities are involved; neurology, cardiology and paediatrics across three different hospitals.

References

[1] Hertzum M and Simonsen J. Effects-Driven IT Development: Specifying and Measuring Usage Effects during Systems Development (forthcoming).

[2] Mller-Jensen J, Pedersen IL and Simonsen J. Measurement of the Clinical Usability of a Configurable EHR. In Hasman A et al. eds. Ubiquity: Technologies for Better Health in Aging Societies, Proceedings of the 20th International Congress of the European Federation for Medical Informatics (MIE 2006), Maastricht, the Netherlands, August 27-30. Amsterdam: IOS Press, 2006; pp. 356-361.

[3] Mller-Jensen J, Simonsen J, and Iversen RK. Measuring Effects on the Clinical Practice from a Configured EHR. In: Hejlesen O et al., eds. Proceedings of the 4th Scandinavian conference on Health Informatics (SHI 2006), Aalborg University, Aalborg, August 24-25. Aalborg: Virtual Centre for Health Informatics, 2006; pp. 58-62.

[4] Simonsen J and Hertzum M. A Regional PD Strategy for EPR Systems: Evidence-Based IT Development. In Jacucci G et al. eds. Expanding Boundaries in Design, Proceedings of the ninth biannual Participatory Design Conference, Vol. II (PDC 2006), Trento, Italy, August 1-5. Palo Alto: CPSR, 2006; pp. 125-128.

Address for correspondence

Anders Barlach
Email: barlach@ruc.dk

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

User Driven, Evidence Based Experimental Design; a New Method for Interface Design Used to Develop an Interface for Clinical Overview of Patient Records

Niels Boye[a,f], Frey Eberholst[b,f], Richard Farlie[c,f], Lene B. Sørensen[d,f], Karen Marie Lyng[e,f]

[a] *Regional Hospital of Randers, University of Aarhus, Denmark*
[B] *Acure – an IBM Division, Aarhus, Denmark*
[c] *Regional Hospital of Herning, Denmark*
[d] *Regional Psychiatric Services, North Jutland , Denmark*
[e] *Tieto-Enator, Denmark*
[f] *Master Education Programme in Health Informatics, University of Aalborg, Denmark*

Abstract

A novel method of interface design - user driven, evidence based experimental design - was developed which approximates the usual clinical way of maturing science and technology in the healthcare domain.

The method is user-driven and the clinician remains in control of gathering and evaluating evidence of relevance to the project - as well as specifying the details of the user interface.

Information not obtainable from the literature was gained experimentally and used to achieve a deeper understanding of the problem before the design phase. The design was subsequently validated experimentally by ordinary users with no connection to the software or design team.

After applying this method to the problem of gaining a satisfactory clinical overview of a single patient's record, we recommend that clinical IT interfaces have clinical logic, sufficient complexity, and are well structured. Developers should use computer power to support "building blocks" such as anatomical problem lists and summaries of history, status and treatment, personal notes, and should support clinical browsing using text and graphics.

Keywords:

medical informatics; computing methodologies; software design; user-computer interface; medical informatics applications; medical record systems, computerized; decisions support systems, clinical; physician practice patterns; clinical reasoning

Introduction

The scientific basis of healthcare is in natural sciences, but the clinical execution also draws on methods and concepts from social sciences. This makes development, formalization, computation and presentation of health related information in electronic systems for clinical use a complex task (1). Attempts to develop a comprehensive and satisfactory overview of a patients status and history to the clinical user in the form of an Electronic Health Record (EHR) has been regarded as inferior to the paper based record.

Various design methods, such as socio-technical design (2-4) and user-centred design or participatory design (www.cpsr.org) have been developed with the aim of promoting a common understanding between IT-professionals and professionals in social sciences such and health care.

In the user-centred design method, the initiative and control in establishing the necessary common understanding between user and IT-professional are in the hands of the latter.

This paper describes the initial efforts to develop a general design method for interfaces, where the common understanding between healthcare- and IT-professionals is established but where control remains with the end users as opposed to the user-centred design method. Furthermore this novel method is more in line with the usual way healthcare technology is matured and developed. We have named it "user driven, evidence based experimental design" to describe the principal components of the method.

The method was developed during work to attempt to solve a general problem in IT systems for health care design; how to obtain a clinical overview of a single patients record. Evidence for the design was drawn from the literature and from our own experiments. The first iteration design was subsequently validated through experiments with end-users (doctors).

Materials and methods

A loosely coupled research group for this specific project was formed at the Master Education programme in Health Informatics at the University of Aalborg, Denmark (http://www.v-chi.dk/english/index.htm). A subgroup of two (RF, KML) investigated "genuine features of overview in paper based records", with a combination of a literature survey and experiments. Another subgroup of two (FE, LBS) investigated "specific features in clinical IT systems conveying the overview", also with a combination of literature survey and experiments. One member (NB) coordinated and supervised the two sub projects. All project members had a common basis of extensive clinical experience, three specialist doctors (RF in gynaecology and obstetrics, KML

in surgery, NB in internal medicine), one nurse (LBS) and one radiographer (FE). No trained IT professionals participated. All participants had years of experience as users and clinical participants in user-design efforts and/or clinical specialists in software companies and all participated in the Master Program in Health Informatics at Alborg University as teacher or students.

Literature overview

Literature surveys and compilation into a common "concept of clinical overview" for the planning and execution of the experimental part was done by Medline searches and subsequent "drilling down" from reviews and papers of cognitive tasks research ((5-9)), decision-theory ((10)), graphical perception ((11-13)) and interface design (e.g. www.useit.com). Literature information was discussed and shared in the total project group.

Empirical information collection

Overview in paper based records

This subgroup employed semi-quantitative interviews with eight doctors from five different medical specialities (four of whom used an EHR in their daily clinical work) to elucidate "the nature of the clinical information overview". This information was analysed using a grounded theory method and led to an observational study with five junior-doctors each performing three constructed scenarios. Data was acquired from the scenarios by the "think aloud method" and various parameters were analysed afterwards from video recordings of each scenario.

Overview in EHRs

A complex patient history based on true data was constructed and presented in two electronic prototypes. The overall construction of the prototypes was similar, but the design in one was mainly text-based and very like EHRs modelled over the paper record, with text and tables.

The other prototype developed was based on the evidence acquired from literature about cognitive mechanisms, clinical and human reasoning, graphics, interface-design and experiments with the overview strategies in paper-based records and it had additional visualisation of clinical information in graphs and displays, constructed in accordance with knowledge of graphic displays and interface design. Both prototypes were programmed "by hand" using Microsoft PowerPoint.

Twenty-three doctors were tested about different problems using either the text-based prototype or the graphics-based prototype of the patient record using the same structured questionnaire. The doctors were randomly allocated to the two groups. User-prototype interaction was observed, timed and further quantified using Camtasia Studio from Techsmith.

Results

Literature review

The literature review showed that doctors use a variety of different strategies to gain an overview of a patient's situation ((14;15)) from the textual information on paper or on screen. The overall strategy employed depends on, whether the information is obtained for the first time or the patient is known to the clinician - as well as time constraints. Reading of textual information can be categorized as: 1) Reading, where the whole text is read, 2) Skimming, where some words in each sentence are read or 3) Skipping, where only few words across a page are read. Hornbæk and Frøkjær (16) have constructed three ways of formatting text: *normal text, fisheyes view* – where less important text-parts were shown with a smaller font, or *overview + detail*, where keywords were extracted and shown in the margin. The *overview + detail view* was found to be the most effective and satisfying for the viewer. A rigorous structure in text and layout and consistent use of design elements supports focused information retrieval (medication, adverse reactions, description of previous results from a procedure) and the general sense of overview by the reader (14). Personal annotations should be supported. Graphical methods should be prioritized and follow the heuristic research based rules for perception of graphical information as set up by Cleveland (11).

Condensation of interview and observational study result for establishing overview of clinical problems using paper based information

Although the paper record employed was unknown to the participants in the observational part, they quickly established an overview of the record structure and formed a preliminary hypothesis for verification or falsification against the specific information in the record by pattern recognition. Data such as lab. results are in tables with a certain structure, daily notes are in ordinary text separated by dates. There was no general agreement on what kind of information it was important to start with.

Observation showed that even junior doctors have developed individual styles in handling the task of establishing an overview of paper based patient record information. Some start from the last entry and read backwards, some start from the beginning and work chronologically towards the most recent entries. The reading style can be combined with ether a tendency to separate the record into several heaps or to keep papers in chronological order, except when comparing individual documents in the record. All alternated between tables, descriptions of results and the entries in the main section of the record. Summaries and compilations of results formulated by colleagues, such as indications for a treatment plan, are easily and quickly identified and used as the "building blocks" of an overview.

Comparison of design elements in the two prototypes

The text based prototype was not just "paper on screen" but contained features that took advantage of computer power to compile chronological overviews, problem lists and lists of active medications and investigations, but maintaining layout and textual features similar to a paper-based record (figure 1). No graphical elements in pathology lab reports, medication lists or diagnostic imaging were present in this prototype.

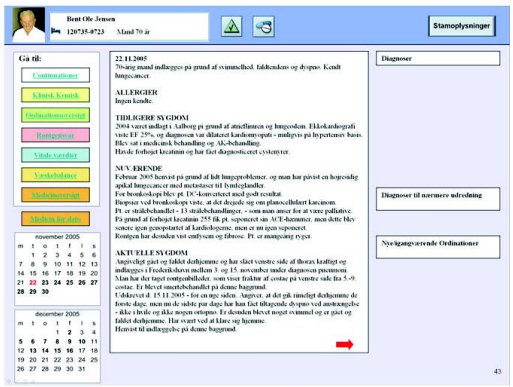


Figure 1 - The text-based prototype with view-menu at the left, a calendar and plain text.

The graphic prototype had the same features as the text based system, with the addition of keyword highlighting to enable text skimming or text scanning, this feature was linked to a graphic anatomical representation of the patients problems. Extensive use of graphical views was used for pathology lab results, vital signs, fluid balance and diagnostic imaging (not suitable for diagnosis) (figure 2).

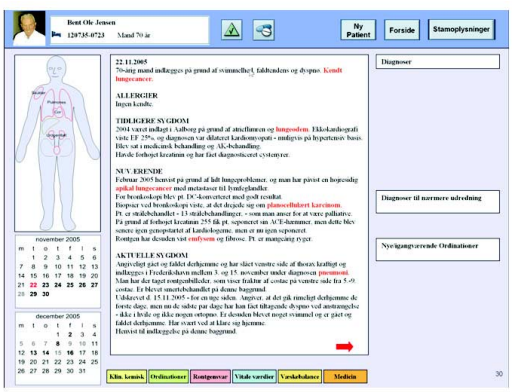


Figure 2 - Same "view" as in figure 1 but in the graphic prototype. The view menu is at the bottom of the screen.

Performance in relation to establishing clinical overview in a text- or a graphic design prototype.

The clinical overview questionnaire of the patient case had 30 questions (each giving 1 point). Ten doctors were allocated to the text based prototype and 13 doctors to the graphic based prototype, but two of these tests partly failed in their quantitative measurements due to technical problems with Camtasia.

Table 1 – Quantitative measurements of performance

Group	N	Mean Time	Range Time	Mean Points	Range Points
Text	10	17:08	14:05-25:00	26.2	23 – 28
Graphic	11	15:38	13:00-17:55	28.1	26 - 30

Time in min:sec

Table 1 shows the quantitative measurements of the two groups solving the clinical questionnaire about the patient. In accordance with the evidence obtained in the project, this table should be displayed graphically to obtain the "non-reduced" picture of the distribution and pattern of performance for each group, but this would demand more space.

The previous IT experience of the doctors in the test was scored by the frequency of IT use, and the number of specific software packages used. The highest score was 22 and the lowest 2. No correlation between IT experience and time used or IT experience and point score in the two questionnaire experiments was found.

The user aspects of the two prototypes were analyzed with a semi-quantitative interview after each test. The text-based prototype was easy to understand, but difficult to navigate, in contrast to the graphics prototype ("you have to get used to this, but it is simple to work with").

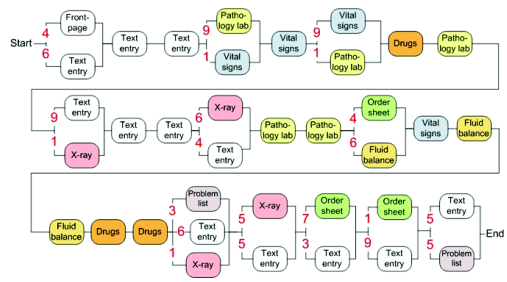


Figure 3 - Flow pattern of the text-based prototype, the number of participants branching is indicated by red numbers.

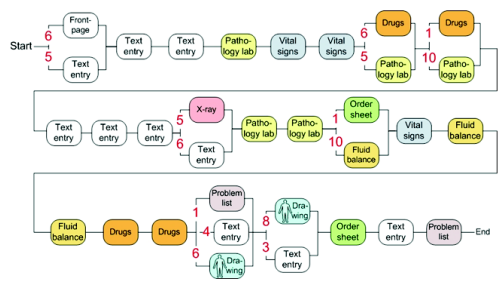


Figure 4 - Flow pattern of the text-based prototype, the number of participants branching is indicated by red numbers

Analysis of the flow and order of screens selected in the text based (10 doctors) and the graphics based (11 doctors) prototype respectively based on data obtained from Camtasia is seen in figures 3 and 4. The graphical prototype gave a simpler and more uniform flow and was generally evaluated higher. Several good ideas for features in future iterations were formulated by the doctors in the interview after the test.

Discussion

The aim of the present project was to develop a method for development of software for clinical use that was closer to the conventional way in which technology is matured, validated, and incorporated in everyday activity in the clinical domain.

Much of the development and subsequent adaptation of technology to clinical surroundings and clinical users in for example, clinical chemistry, molecular biology and diagnostic imaging has been done endogenously in the clinical organization, based on scientific (in contrast to commercial) sharing of knowledge, and furthermore driven by experimental methods. This implies that the clinicians are in control of the experiments and validation of results, with the scientific "responsibility" to colleagues and peer-review bodies. It will still be a cross disciplinary effort to develop clinical information systems and maturation of concepts to an operational level, as development of sufficiently advanced software platforms for hospital use will probably still be a commercial activity.

The general clinical problem used as a "case" in this paper was the overview of previous documentation in a patient's record in order to obtain a satisfactory understanding of the patients' history and present status, as the basis for new clinical (human-driven) reasoning. It can, of course, be questioned whether this can be mimicked by the ability to answer 30 specific questions about the patient using two different electronic interfaces. More experiments with more patients, cases and interfaces are needed to elucidate this aspect.

Using Microsoft Power Point as the tool for visualization of concepts is not advisable, since it becomes unpredictable when working with complex presentations containing many internal hyperlinks, but we used it because we could not find a better software package for the purpose at the time.

The validation experiments are crucial to the method, since it separates the developers from the feed back from end users in a sincere way, which, although it is cumbersome, will give more convincing results for use in the next iteration. It is of great importance to use structured and methodologically sound validation experiments to obtain evidence of suitable strength,

The graphics based prototype gave the highest score and the simplest flow of the two prototypes with the same information content. No statistical methods have been employed, since the differences are more qualitative than quantitative. The authors interpretation is, that the graphic prototype to a larger extend supported the users need to alternate between "the overview" and "the detail" without losing track when investigating a clinical problem. This is probably also the mechanism behind the more effective text reading of skimming and scanning and the *overview+detail view* of Hornbæk and Frøkjær (16).

On the basis of our literature studies and experiments, we have the following recommendations for interface design for clinical use. This advice will need further validation in future projects and experiments.

> Medicine is a visual, 3-dimensional discipline. This should be reflected in the interface.
>
> A rigorous, clinically intuitive structure in text and graphic displays, and the use of design elements with easily recognizable patterns is advised. The interface does not have to be simple, since the clinical user can quickly master a complicated design if it is "clinically logical", and there is no need to strive for a standard program layout – clinical functionality is more important.
>
> Browseability in text should be supported by different text presentations such as standard text, fisheye and overview + detail. Automatic highlighting of clinical terms can increase browseability. Filters for different types of entries could increase the acquisition of overview "building blocks".
>
> Automatic generation of problem lists, active medication and vital signs charts etc., and an anatomical view of the problem list (and other information) were rated as desirable features in our test group. Computer power should be used to ease the cognitive burden of the clinical user and will probably give a good "return on investment" (17). Opportunities for personal annotations and search facilities should be provided. A list of previous cases for junior doctors with notification facilities would aid personal reflection on performance and learning.
>
> Some indication of the complexity of the clinical problem and summaries of previous information should be easily displayed.

The browseability features of the paper record such as scattering documents over a desk top can only be partly recreated in the EHR by employing large or multiple displays. Gesture control of objects and object relations on a large display may be desirable in terms of "computer supported augmented clinical overview."

References

[1] Kanstrup A, Boye N, Nohr C. User-driven design considerations of m-learning services for residents - activation of a theoretical model of clinical knowledge. Medinfo2007, Brisbane 2007.

[2] Berg M. Working with protocols: a sociological view. Neth J Med 1996; 49(3):119-125.

[3] Berg M. Medical work and the computer-based patient record: a sociological perspective. Methods Inf Med 1998; 37(3):294-301.

[4] Berg M. Patient care information systems and health care work: a sociotechnical approach. Int J Med Inf 1999; 55(2):87-101.

[5] Evans D, Patel VL. Advanced Models of Cognition for Medical Training and Practice. Elsivier, 1991.
[6] Patel VL, Kaufman DR, Arocha JA, Kushniruk AW. Bridging theory and practice: cognitive science and medical informatics. Medinfo 1995; 8 Pt 2:1278-1282.
[7] Kushniruk AW, Patel VL. Cognitive computer-based video analysis: its application in assessing the usability of medical systems. Medinfo 1995; 8 Pt 2:1566-1569.
[8] Patel VL, Kushniruk AW. Interface design for health care environments: the role of cognitive science. Proc AMIA Symp 1998;29-37.
[9] Kushniruk AW, Kan MY, McKeown K, Klavans J, Jordan D, LaFlamme M et al. Usability evaluation of an experimental text summarization system and three search engines: implications for the reengineering of health care interfaces. Proc AMIA Symp 2002;420-424.
[10] Wulff H, Gøtzsche P. Rationel Klinik. 4 ed. Copenhagen: Munksgaard, 1997.
[11] Cleveland WS, McGill R. Graphical Perception and Graphical Methods for Analyzing Scientific Data. Science 1985; 229:828-833.
[12] Cleveland WS. Visualizing Data. Hobart Press, Summit, New Jersey, 1993.
[13] Cleveland WS. The Elements of Graphing Data. Revised edition ed. Hobart Press, Summi, New Jersey, 1994.
[14] Nygren E, Henriksson P. Reading the medical record. I. Analysis of physicians' ways of reading the medical record. Comput Methods Programs Biomed 1992; 39(1-2):1-12.
[15] Nygren E, Wyatt JC, Wright P. Helping clinicians to find data and avoid delays. Lancet 1998; 352(9138):1462-1466.
[16] Hornbæk K, Frøkjær E. Reading patterns and usability in visualizations of electronic documents. ACM Transactions on Computer-Human Interaction (TOCHI) 10 (2)[June 2003], 119-149. 2006. ACM press New York, USA. Ref Type: Electronic Citation
[17] Kaushal R, Jha AK, Franz C, Glaser J, Shetty KD, Jaggi T et al. Return on investment for a computerized physician order entry system. J Am Med Inform Assoc 2006; 13(3): 261-266.

Address for correspondence

Niels Boye, Regional Hospital of Randers,
Skovlyvej 2,
DK-8900 Randers
Denmark
E-mail: Niels@Boye.dk

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

User Interface Optimization for an Electronic Medical Record System

Kai Zheng[a], Rema Padman[b], Michael P. Johnson[b]

[a] *School of Public Health and School of Information, University of Michigan, U.S.A.*
[b] *H. John Heinz III School of Public Policy and Management, Carnegie Mellon University, U.S.A.*

Abstract

Many information technology-enabled healthcare applications have failed because their interfaces are difficult to use. Unfortunately, little attention has been paid in the health informatics community to designing effective user interfaces that are acceptable to healthcare professionals. This paper illustrates a method for improving application interface usability by applying sequential pattern analysis to analyze temporal event sequences recorded in an electronic medical record system. Such event sequences, or clickstreams, reflect clinicians' navigation patterns in their everyday interactions with the computer system. The identified patterns have been used by software developers to calibrate the user interface of the system, so that the within-application workflow is better aligned with clinicians' mental model of medical problem-solving. Such inferred patterns may also help to modify clinicians' suboptimal practice behavior components, as manifested through their actual usage of this point-of-care electronic system.

Keywords:

user-centered design; user interface design; sequential pattern analysis; human-computer interaction; usability assessment; data display

Introduction

Medical practice is a complex process. Large amount of data must be accessed, assembled, and analyzed at the point of care to inform proper medical decision-making. In the era of paper-based patient records, clinicians flip through stacks of paper charts to look for desired information. The use of electronic systems has greatly facilitated health data retrieval. However, it has also introduced new dimensions of problems. Two paper documents, for instance, can be laid out side by side for cross reference, while on a computer screen it is usually impractical to have two windows visible at the same time. How to preserve the easy "look-and-feel" of paper charts is a real challenge for software developers. In addition, poorly designed application navigation flow may also escalate learning effort, decrease productivity, and increase user errors [1, 2].

Lack of good user interfaces has been long recognized as a major impediment to the acceptance and routine use of clinical informatics applications [3]. Unfortunately, very few research studies have looked at design principles for building intuitive and effective healthcare user interfaces (UI); even fewer have validated the usability of existing UI design in realistic clinical settings. Consequently, "systems are created *ad hoc*, users are dissatisfied, and often systems are abandoned" [2].

The present study was motivated by these facts. The computer system in question, the *Clinical Reminder System* (CRS), is a "*lite*" electronic medical record system (EMR) that collects, stores, and manages a wide range of patient and clinical data [4, 5]. In addition to its regular EMR functionalities, CRS is also intended to improve quality of care by providing clinicians "just-in-time" alerts and advisories using evidence-based guidelines.

Since 2002, CRS has been deployed in an outpatient clinic at an urban hospital, and used by clinicians to treat patients in real time. While user, task, and representational analysis were performed during the software design phase with constant feedback by participating clinicians, its UI design was still critiqued after being routinely used in clinicians' everyday practice. In a user satisfaction survey following a 10-month field trial, users complained that the application's early user interface, shown in Figure 1, provided little guidance as to a desired workflow [4, 5]. As a result, user acceptance was not satisfactory, and the utilization rate of the system remained low [4, 5].

Although this UI reflected the best knowledge of developers and preferences of the client organization, the standard Windows-based layout was reported as "not aligned with our common practice styles". The horizontally arranged tabs, for example, did not reflect the preferred order of clinical information access. As a result, users expended substantial energy unnecessarily to adapt their practice to a UI design that they considered "uncomfortable".

Selected for best paper award.

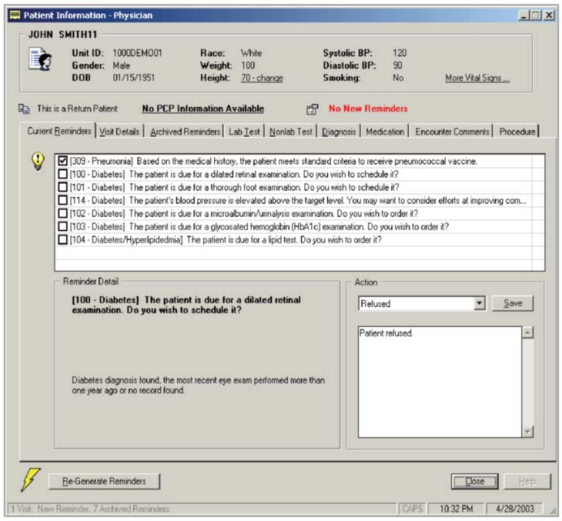


Figure 1 - An early user interface

To solve the identified UI flaws, the system was reengineered into a full web-based application. A screenshot of the new web interface is shown in Figure 2. Unique features of the web-enabled application provide tremendous promise for maximally preserving the "look-and-feel" of traditional paper charts. In the new design, for example, different features conveying different clinical information elements are no longer arranged in a tabular form, instead, they are displayed in the same workspace that can be easily navigated by mouse scroll wheels, simulating paper-flipping behavior. A navigation menu is also provided on an adjacent frame to enable fast switches across different features. However, it is not known whether this new design is consistent with, or represents an improvement upon, clinicians' typical workflow.

The study reported in this paper was therefore conducted to identify the preferred sequential order in which different features of the system are accessed. To learn clinicians' navigation behavior, this study uses a sequential pattern analysis method to analyze actual usage recorded in the computer logs that contain time stamped events. Actual usage data, unlike many software usability experiments, represent users' interaction with a system under real working conditions, rather than on contrived laboratory exercises.

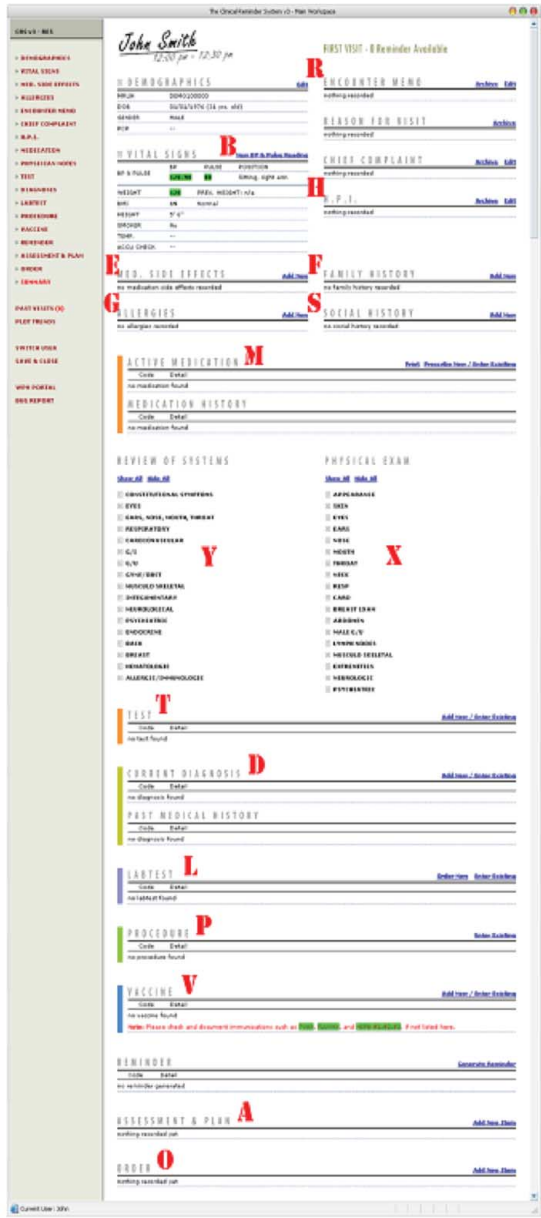


Figure 2 - New user interface to be evaluated

Methods

Sequential pattern analysis

Sequential pattern analysis discovers hidden and recurring patterns within large sequences of events. It has been applied in a wide variety of domains such as web person specialization and page recommendation [6], HCI usability testing [7], and genetic sequence analysis [8]. In this study, a consecutive sequential pattern algorithm is employed to analyze the event sequences recorded in CRS. This algorithm detects consecutively occurring events that appear across different sessions. Such patterns, that represent adjacent feature accesses frequently occurring next to each other and in a given sequential order, are of particular interest to inform UI redesign.

Let s denote an even sequence by $<e_1, e_2, \ldots, e_n>$, where $e_j, j \in 1 \ldots n$, is an event that occurs at the j^{th} position in s. The consecutive sequential pattern algorithm finds a sequence $p <p_t, p_{t+1}, \ldots, p_{t+l}>$ that is a subset of s, which is also part of, or supported by, other sequences. The support for p is defined as the fraction of total sequences that supports p. When a sequence satisfies a certain minimum

support threshold, it is named a *Sequential Pattern*. The largest length sequential pattern that is not part of any other patterns is called a *Maximal Sequential Pattern*. The objective of the sequential pattern analysis is to find all such maximal sequential patterns.

When the minimum support is a constant for any given length, the most efficient algorithm starts with calculating support for all possible sequences composed of two consecutive events. When a sequence does not satisfy the minimum support, it is removed from further computation; otherwise, it is treated as a candidate sequence to compute support for larger length sequences. The algorithm stops when no larger length sequences based on a current candidate would satisfy the minimum support. The current candidate sequence is then chosen as a maximal sequential pattern.

Study site and data collection

In this study, 10 months of usage data were electronically collected from October 1, 2005 to August 1, 2006 and analyzed. These usage data were generated from the most recent web-enabled version of CRS. The system implementation was accomplished in the summer of 2005 and substantial training was provided afterwards.

The main CRS user population during the study period was composed of 40 first-, second-, and third year internal medicine residents. Residents who used the system for fewer than 5 patient encounters are excluded from the analysis. It is likely that such users' interactions with CRS do not reflect mature application usage. 30 active resident users were thus identified, whose system usage was recorded in 973 unique patient encounters.

Data analysis and results

Data preparation

Data preparation procedures were performed prior to the analysis. All events and their affiliated attributes, such as session ID and time stamp, were first collected from scattered data tables. The event type was then mapped based on a labeling schema, which is composed of distinct letter symbols. Table 1 lists all 17 main features[1] that the CRS application provides, ordered alphabetically by their labeling symbols[2]. The screenshot shown in Figure 2 illustrates the on-screen positions of each of the 17 major features.

Event sequences were then constructed. `HMMMYAD`, for instance, is a 7-length sequence composed of 7 events that occurred within a patient encounter, ordered chronologically by their time stamps. The resulting event sequences are further consolidated by collapsing repeating access to a same feature. For example the segment `MMM`, "prescribing multiple medications consecutively", is collapsed into one single event `M`. In this study only across-feature navigation is of interest, that is, "jumps" across different features.

Figure 3 shows the distribution of event sequence length after the collapsing operation. The sequences composed of 4 or less events are excluded from further data analysis because they provide little information in regard to sequential navigational patterns. This operation results in the loss of 6 additional users whose recorded sequence lengths are all below 5. After these data preparation procedures 473 event sequences are retained, generated by 24 distinct resident users. Distribution of number of sequences owned by each user is depicted in Figure 4. Several sample event sequences are shown below:

```
HMOMYXAM
GHXVHADADHA
HGYXADAOMYSX
OMRHFYXYXADADA
HXOPMOMOMOMODADAM
HSXDADADADADAMOMOMOMOMO
```

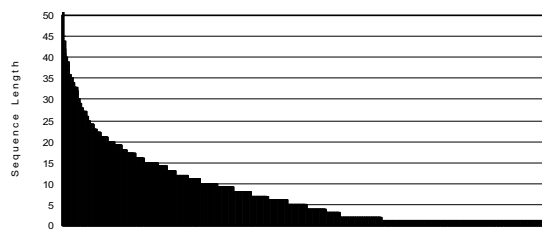


Figure 3 - Distribution of event sequence length

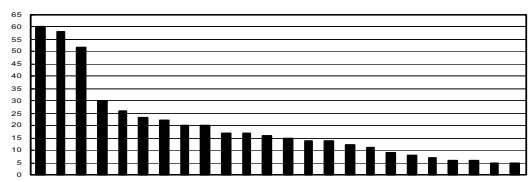


Figure 4 - Usage distribution among users

Frequency of feature access

Table 1 shows the aggregated proportion of feature accesses[3]. These proportions roughly represent how frequently each application feature was used. As shown in Table 1, among the 17 major features "Assessment and Plan", "Diagnosis", and "Medication" were most heavily used. Note that while "Encounter Memo" appears on top of screen, it was seldom accessed.

1 Feature that must be displayed in a certain position for legal reasons, such as patient's demographics always appearing on top of an encounter page, is excluded from the consideration of this study. Also excluded are non-actable or not yet activated features, for example "Reason for the Visit" that is entered by nurses when a patient calls to make an appointment.

2 A symbol letter is usually the first letter of a feature unless there is a conflict.

3 Repeating access to a same feature is counted only once.

Table 1 - Main features and overall frequency of access

Label	Feature	Proportion (%)
A	Assessment and Plan	21.18
B	Retaking BP	.34
D	Diagnosis	16.36
E	Medication Side Effects	.22
F	Family History	1.24
G	Allergies	1.88
H	History of Present Illness	7.26
L	Laboratory Test	3.58
M	Medication	14.53
O	Order	17.17
P	Procedure	.38
R	Encounter Memo	.44
S	Social History	2.85
T	Office Test	.62
V	Vaccine	.83
X	Physical Examination	6.69
Y	Review of Systems	4.43

Table 2 shows the results of the sequential pattern analysis. All maximal sequential patterns included in the table satisfy a minimum support threshold of 15%. These patterns are sorted by the level of support they received.

Table 2 - Maximal sequential pattern discovered

Maximal Sequential Pattern	Level of Support (%)
ADAD	51.16
DADA	43.97
XADA	40.17
OMOM	32.77
MOMO	29.39
YXAD	21.78
HS	19.03
OL	18.6
OMY	16.7
LO	15.64
HO	15.01

Some interesting sequential patterns emerge from Table 2. ADAD, appearing in 51.16% of all encounters, is the most salient pattern discovered, followed by a similar and partially overlapped pattern DADA, with 43.97% support. It indicates that the users of CRS frequently switched between the features "Assessment and Plan" and "Diagnosis". Similarly, users frequently switched between "Order" and "Medication", with 35.1% support for OMOM and 27.06% support for MOMO; and "Order" and "Laboratory Test", with 18.6% support for OMOM and 15.64% support for MOMO. A further examination found that A precedes D more often (89.18%) when a user entered the AD...AD or DA...DA segment. Similarly, O was usually accessed before M (72.57%), and before L (71.58%).

Supported by 40.17% of all encounters, ADA is preceded by X - "Physical Examination", and YXAD appears in 21.78% of time. This indicates that "Physical Examination" "Assessment and Plan" "Diagnosis" is a frequently traversed path, which is often preceded by accessing "Review of Systems". Further, OMY occurs in 16.7% of all sequences, indicating that OM - "Order" and "Medication" were often used before Y - "Review of Systems". HS - "History of Present Illness" then "Social History" and HS - "History of Present Illness" then "Order", are two other consecutive patterns with slightly smaller support, 19.03% and 15.01%, respectively.

An *ad hoc* within-sequence analysis was further conducted to detect sequence segment recurring within an encounter session. Results are shown in Table 3. The "Probability of Repeat" in Table 3 exhibits the probability of a two-length event segment recurring within a sequence. DA or AD - "Diagnosis" and "Assessment and Plan", OM or MO - "Order" and "Medication", and OL or LO - "Order" and "Laboratory Test", are three frequently repeating segments thus identified, which also confirm the cross sequence patterns of DADA, ADAD, OMOM, MOMO, OLOL, and LOLO. Because items in these reappearing sequence segments were usually accessed next to each other, they are hereby referred to as *Bundled Action.*

Table 3 - Recurring patterns within encounters

Sequential Pattern	Probability of Repeat (%)
AD	70.22
MO	64.98
OL	64.77
DA	64.35
OM	63.67
LO	51.35

The repeating access to bundled actions, however, blurs the boundary of "jumps" from a series of bundled action accesses to other features. For example the reappearing AD with varying length in the sequence HADAD...ADADXY impairs the analytical power for discovering whether there exists a pattern H-AD-Y that may help reveal interesting patterns at an overall level. Similar to collapsing repeating access to the same feature, repeating access to the same bundled action is further collapsed to count as one single occurrence. For example the HDAD...ADADAXY sequence is converted into HDAY to form a new, higher level sequence.

A second pass sequential pattern analysis was then conducted to analyze the event sequences obtained after this collapsing operation. ADO - "Assessment and Plan" to "Diagnosis" to "Order" is the only additional sequential pattern thus identified, supported by 15.64% of all encounters. This pattern indicates that after a user finished working on "Assessment and Plan" and "Diagnosis", he or she would switch to the "Order" section immediately to prescribe orders of new medications or laboratory tests.

Discussion

Based on the findings from analyzing actual usage data with sequential pattern analysis, several UI design principles can be arrived at:

- "Encounter Memo" should be properly relocated. This feature is less frequently used while occupying the most salient position in the current design;

- "Assessment and Plan", "Diagnosis", and "Medication" are the most frequently accessed features. They should be placed in the most salient positions on a computer screen;
- "Assessment and Plan" and "Diagnosis", "Order" and "Medication", and "Order" and "Laboratory Test" are *bundled actions*. They are usually accessed next to each other and often used multiple times within an encounter session. Navigation aids such as hyperlink shortcuts should be provided to facilitate these frequent feature switches;
- "Review of Systems", "Physician Examination", "Assessment and Plan", and "Diagnosis" should be presented adjacent to each other in this sequential order. Accesses to these four features often appear as a series of events occurring sequentially.

These design principles have been used in redesigning the existing user interface of CRS. Since the basic EMR functionalities that CRS provides are universal, these design principles may also be applicable to other electronic medical record systems.

Conclusions

Improving the UI design of an electronic medical record system can be successfully attained by analyzing the actual usage data recorded during its everyday use. The sequential patterns identified in this paper led to a set of design principles used in redesigning the application's user interface. These design principles mainly propose that different clinical information elements should be presented in the sequential order in which they are usually accessed, which reflects clinicians' mental model of medical problem-solving during patient encounters.

This study has a few limitations. First, actual usage data must be collected from a working system. Its current design, inevitably, may exert an influence on users' own working style. Second, the findings are derived from testing a single system with certain unique features. While the method and the results provide general insights into designing user interfaces for other types of health applications, they may not be used without careful customization. Finally, the user population of this study was mainly composed of internal medicine residents. The derived design pattern reflecting their practice style may not be generalizable to other clinical specialties.

References

[1] Bates DW, Kuperman GJ, Rittenberg E, Teich JM, Fiskio J, Ma'luf N, Onderdonk A, Wybenga D, Winkelman J, Brennan TA, Komaroff AL, and Tanasijevic M. A randomized trial of a computer-based intervention to reduce utilization of redundant laboratory tests. Am J Med 1999; 106(2):144-50

[2] Tang PC and Patel VL. Major issues in user interface design for health professional workstations: summary and recommendations. Int J Biomed Comput 1994; 34(1-4):139-48.

[3] Johnson CM, Johnson TR, and Zhang J. A user-centered framework for redesigning health care interfaces. J Biomed Inform 2005; 38(1): 75-87.

[4] Zheng K, Padman R, Johnson MP, Engberg J, and Diamond HS. An adoption study of a clinical reminder system in ambulatory care using a developmental trajectory approach. Medinfo. 2004; 11(Pt 2):1115-9.

[5] Zheng K, Padman R, Johnson MP, and Diamond HS. Understanding technology adoption in clinical care: clinician adoption behavior of a point-of-care reminder system. Int J Med Inform 2005; 74(7-8):535-43.

[6] Montgomery AL, Li SB, Srinivasan K, and Liechty JC. Modeling online browsing and path analysis using clickstream data. Marketing Science 2004; 23(4):579-95.

[7] Siochi AC and Ehrich RW. Computer analysis of user interfaces based on repetition in transcripts of user sessions, ACM Transactions on Information Systems 1991; 9(4):309-35.

[8] Eres R, Landau GM, and Parida L. Permutation pattern discovery in biosequences. J Comput Biol 2004; 11(6):1050-60

Address for correspondence

Kai Zheng, kzheng@umich.edu, M3531 SPH II, 109 S. Observatory, Ann Arbor, 48109-2029, U.S.A.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

AdaRTE: Adaptable Dialogue Architecture and Runtime Engine. A New Architecture for Health-Care Dialogue Systems

L. M. Rojas-Barahona[a], T. Giorgino[a]

[a] *Laboratory for Biomedical Informatics, Dipartimento di Informatica e Sistemistica, Università di Pavia, Italy*

Abstract

Spoken dialogue systems have been increasingly employed to provide ubiquitous automated access via telephone to information and services for the non-Internet-connected public. In the health care context, dialogue systems have been successfully applied. Nevertheless, speech-based technology is not easy to implement because it requires a considerable development investment. The advent of VoiceXML for voice applications contributed to reduce the proliferation of incompatible dialogue interpreters, but introduced new complexity. As a response to these issues, we designed an architecture for dialogue representation and interpretation, AdaRTE, which allows developers to layout dialogue interactions through a high level formalism that offers both declarative and procedural features. AdaRTE aim is to provide a ground for deploying complex and adaptable dialogues whilst allows the experimentation and incremental adoption of innovative speech technologies. It provides the dynamic behavior of Augmented Transition Networks and enables the generation of different backends formats such as VoiceXML. It is especially targeted to the health care context, where a framework for easy dialogue deployment could reduce the barrier for a more widespread adoption of dialogue systems.

Keywords:

telemedicine, speech recognition software, ambulatory care information systems, telephone, computerized, chronic obstructive pulmonary disease, hypertension

Introduction

Dialogue technologies have been proven useful to provide the general public with access to telemedicine services. Several studies have discussed their advantages for chronic symptoms monitoring, interviews, counseling, education, etc. [1, 2]. Dialogue systems in health care context are deployed to complement traditional contact channels and have been used for several home-care interventions successfully [3-7]. Thus, they may be able to improve quality of service and communication in a cost-effective way.

In previous projects, customized technology was the response to vocal applications deployment. Available technology allowed implementation either via custom code, or proprietary dialogue-manager based solutions. Several dialogue manager based architectures have been devised in order to simplify complex programming present in custom-coded applications. In addition, the multitude of dialogue technology vendors naturally resulted in a proliferation of incompatible languages across vendors and platforms.

Recently, the concept of Voice Browser (VB) was introduced by W3C [8]. VBs foresee a dialogue-manager, which understand VoiceXML documents instead of a proprietary language. Despite providing a great deal of independence from speech recognition engine vendors, VoiceXML has serious shortcomings in allowing the reuse of components, database access and support of natural language processing (NLP) and multimodality [9]. Furthermore, the sequence of dialogue steps in VoiceXML is defined with a sort of form-filling mechanism and, as web based technologies, has to be generated dynamically by other code. In general, visual inspection and maintenance of scripts is less than straightforward.

The need for a leaner development methodology is especially evident when considering domains in which the manpower and the time-cost available for development are limited, such as health domain. In this paper, we present a novel architecture, AdaRTE devised in order to overcome the issues of the existing dialogue management methods. AdaRTE features were thought to reduce dialogue system development effort through re-use, support of augmented transition networks, adaptable decision takers and best practices adoption. We built AdaRTE, which implements these features for dialogue deployment, and we present the results obtained through the partial prototyping of two telephony-linked systems: the first inspired by the Chronic Obstructive Pulmonary Disease (COPD) care [5], and the second by the Homey dialogue system for hypertensive patient home management [6].

Our effort was mainly focused on health care dialogues systems, since our solution is especially targeted at offering low cost, standards-compliant deployment and experimentation through the incremental integration of other voice formalisms i.e. NLP based on lexicalized grammars.

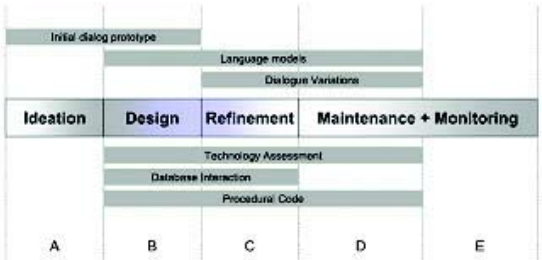


Figure 1 - Phases in the Homey development cycle

Background

A range of technologies are available for building dialogue systems. The simplest of these technologies is a linear script; others are state transition networks and plan-based dialogue systems [7]. Generally, the deployment of any of these techniques requires heavily scripted solutions. Additionally, there is not much information available about time and costs implied in dialogue systems development process. As a matter of fact, deployment of dialogue systems was considered art rather than engineering or science, because of the scarcity of standards. Commonly, the process of deploying dialogue systems was complicated, costly, time demanding and required speech technology experts. To give a specific example, in one previous projects, the EU-sponsored project Homey [6], the time spent in developing the technical aspects of the system has been rather long and the result was reusable only to a limited extent (Figure 1). The voice part of the Homey system required approximately one man-year worth for design and implementation. The Maintenance and Monitoring milestone targeted issues roughly grouped in the following areas: Incorporation of procedural code, database access, language models, increased variability of speech, prompts, modularization and issues related to ASR engine dependency.

More recently, on the other hand, the maturation of VoiceXML allows to deploy dialogue systems in a Web-based environment [8]. Its delivery contributed to reduce the proliferation of incompatible dialogue formalisms by offering one standard for voice applications, so that platform independence and simplicity of development are allow at the same time. Several dialogue systems deployed by using VoiceXML have been published lately. For instance, a dialogue for diabetes home monitoring was implemented by integrating VoiceXML with a Voice Service Provider (VSP), developing a visual user interface and the database backend in [10].

In spite of the advantages explained above, VoiceXML has inherent limitations which are well analyzed in [9, 11], such as its declarative and static structure, difficulty accessing remote resources (databases and ontologies) and lack of means for efficient and heavy computation. Furthermore, the strongest limit pointed out by the research community is that neither dynamic natural language understanding and generation nor multimodality is directly supported. As a consequence, a variety of extensions to VoiceXML has been proposed: for instance, DialogXML was applied to car telematics, in this approach the VB was extended to support NLP KANTOO generated grammars [12]. A prototype of an editor for creating VoiceXML documents is exposed in [13]. Despite the emerging VoiceXML-generative frameworks, we believe that a big effort should still be done in adapting dialogue systems best practices such as confirmation strategy, adaptability, mixed initiative, usable speech interfaces for users and graphical interfaces for developers, together with innovative speech solutions including NLP in VoiceXML based frameworks. [14]

Methods

Our proposed architecture, shown in figure 2, is primarily composed of a *dialogue interpreter, a runtime engine* and *an interface media realizer for backends generation.* A running system interacts with users which can be grouped in three main role categories: Application developers, patients and case managers, i.e. case manager nurses.

In order to enable rapid prototyping, every given dialogue should be developed in a graphical environment which will allow the layout of prompts, speech items, and updating their properties. The editor will represent and store the dialogue structure in a well-defined formalism. To simplify notation, this representation will also be called XML dialogue description from now on.

To cooperate with standards-based speech recognition software and respond to telephone-originated events, AdaRTE acts as a web server, dynamically generating VoiceXML code. The code is sent to the VB over HTTP and a local network connection. The VB, in turn connected to telephony hardware, will be in charge of interpreting documents generated according to user's interaction over the phone. The browser captures and recognizes the answers, and streams them back to AdaRTE through an HTTP *post* request.

Prompts, questions and other elements are the nodes (here named *blocks*) of an Augmented Transition Network (ATN) that specifies the flow of the conversation. Blocks, shown graphically in Figure 3, are represented in the description by XML tags. When the system is started, the XML dialogue description is read by AdaRTE which maintains an internal representation of the dialogue, and executes it when a call comes in. Consequently, it activates the dialogue blocks in sequence or according to a specific

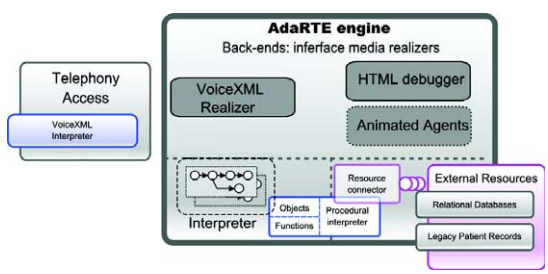


Figure 2 - AdaRTE architecture block diagram

criterion, constructs prompts, interprets the answers returned by the caller through the voice platform, and interacts with external resources as appropriate.

Usually, an ATN is associated with a context or a topic and here we call this structure *subdialogue*. The purpose of subdialogues is to help partitioning a complex application into modules. This contributes to structure the conversation layout for ease of maintenance by allowing reusing subdialogues within the same application, and makes it easy to reuse dialogue component blocks between applications (figure 4).

When a call is setup, the main subdialogue is retrieved and started; it can in its turn invoke other subdialogues, and so forth. If the execution flow reaches the end of the main subdialogue, the call is terminated. Subdialogues can also terminate unexpectedly if an exception occurs, and an exception handler is executed.

Blocks available for nesting within subdialogues are: prompt, question, script, decision, exception handler, promptset, placeholders, containers and subdialogue calling blocks.

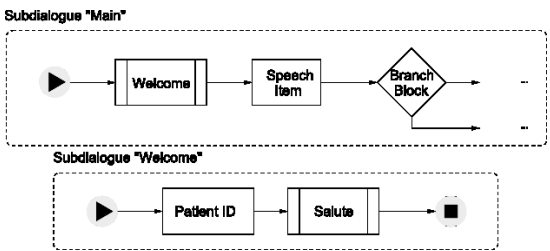


Figure 3 - Block-based dialogue description. Subdialogues are defined by the application developer (shown here as rounded dotted boxes), and can be invoked with a subdialogue calling block (shown in double-border)

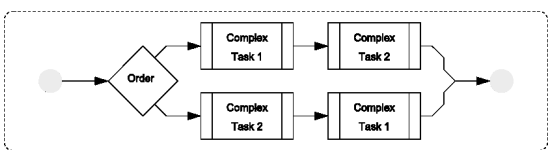


Figure 4 - Within-application reuse of blocks. Dialogue sequence can be rearranged without block duplication.

In addition, we grant the configuration of skip lists for the confirmation strategy related to questions. Containers permit designers to implement adaptability in the dialogue i.e. flexibility according to users experience with the system. Containers are used for common tasks in which one of several subdialogues is selected according to a specified policy (figure 5). Policies for activations of blocks inside containers could be: randomly, in sequence, ordered by call number and according to an externally defined schedule. Another policy could be generated by performing statistical tests to classify the level of experience of a user or even the likelihood of his/her encountering problems on specific parts of the dialogue.

Inclusion of procedural code is essential for flexibility, interoperability, and ease of programming. AdaRTE allows embedding snippets of code written in the ECMAScript standard language, into *script blocks*. The user-written code is run in a separate execution environment with extensive facilities and standard libraries. This also enables external resources access such as databases, ontologies, or any other commodity library.

Currently, the semantic recognition is implemented inside each question, and support the context-free grammar (CFG) formats offered by the VB [15]. However, we are working in the integration of a more elaborated semantic recognition solution by supporting NLP and lexicalized grammars which are more expressive than CFG formats.

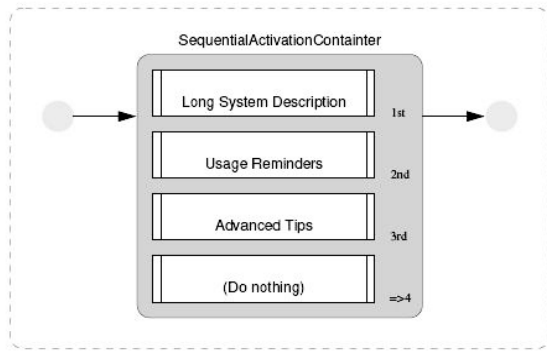


Figure 5 - Containers automate switching between homologous blocks. Switching happens according to a container-specific policy — in this case, only one contained block is activated per invocation, according to call number for that patient: a long system description is played on the first call, a brief reminder given on the second time he calls, and so on. Containers simplify the addition of variability to the dialogue

Results and discussion

The AdaRTE framework is currently in operation. It has been beta-tested with two realistic health care dialogue systems, derived by actual systems deployed and validated in the previous years. The first one is based on a prototype based on the TLC-COPD dialogue deployed in the past by the Boston MISU group and others [5]. For this specific example, we used Tellme Studio[1] as VSP. This pilot's deployment demanded less than two weeks of man effort. The fulfilled activities such as database schema definition and data preparation together with dialogue deployment are shown in figure 6a. This dialogue is executed in English language and uses keypad touch-tone (DTMF) interaction.

The second test case is the partial re-implementation of the Homey dialogue system. Homey had been deployed and evaluated in two Italian hospitals for the management of

1 Tellme Studio. https://studio.tellme.com/

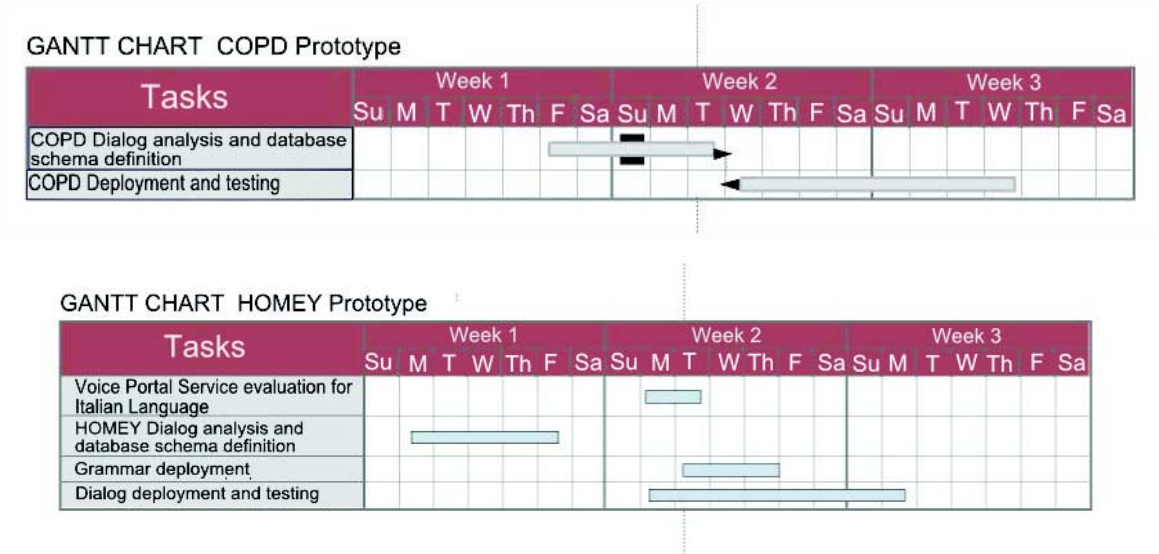


Figure 6 (a) Gantt diagram of the COPD dialogue pilot prototype.
(b) Gantt diagram of the "Hypertension" pilot deployment.

hypertensive patients [6]. The system included an extensive Electronic Health Records system with storage of personal data and profiles, in order to support dialogue adaptivity. Re-engineering the system from the original proprietary dialogue manager to the AdaRTE architecture took approximately three weeks (eleven days of man effort). The development of this prototype involved the following activities: VSP evaluation, database definition and grammars and dialogue deployment (figure 6b). Unlike the TLC-COPD pilot, this system uses *speech* rather than DTMF input. We built grammars using the Nuance GSL language (Nuance 7) and SRGS grammar formats [15]. The language of the dialogue is Italian and the dialogue was deployed by using Voxpilot as VSP[2].

The expressiveness of the dialogue formalism yielded an important reduction of the time invested in developing these two prototypes. Examples of actual blocks' implementations are detailed in figures 7 and 8. Figure 7 shows the main subdialogue of the COPD partial implementation; an example of a script block which embeds a function that retrieve some patient's store data is shown in figure 8.

AdaRTE differs from other similar frameworks in that it is targeted towards the medical domain, which requires adaptable dialogs with complex structures and enquiry data collection tasks. Also, it offers a new level of flexibility to developers by allowing external resources access through the procedural features implemented inside script blocks, at the same time simple dialogs could be implemented by not expert authors. Finally, AdaRTE was thought to be a standard-compliant extensive architecture for the incremental adoption and experimentation of innovative speech technologies formalisms.

```
<script id ="2" next="3">
 <![CDATA[
          initalize_vars();
          function setPatientCode(pin){
                    var conn = DBConnection.getConnection("iper");
                    var statement = conn.createStatement();
                    var resultSet = statement.executeQuery("SELECT nopaziente,
                              CONCAT(CONCAT(nome,' '),cognome) AS nomePaziente
                              FROM anagrafica WHERE cod_telefono like '" + pin +  "' " );

                    while (resultSet.next()) {
                              nopaziente = resultSet.getInt("nopaziente");
                              nomePaziente = resultSet.getString("nomePaziente");
                    }
                    ....
 ] ]>
</script>
```

Figure 7 - Top-level dialog sequence (COPD example)

```
<subdef name="main">
   <start id="1" next="2"/>
        <subdialog id="2" next="3" name="identification"/>
        <if id="3" next="4" cond="followUpCALL == '1'">
              <else next="5"/>
        </if>
        <subdialog id="4" next="6" name="FollowUpCall"/>
        <subdialog id="5" next="6" name="dyspnea"/>
        <subdialog id="6" next="7" name="closingStatement"/>
        <catch event="dialog.finishCall" next="7">
            <start id ="1" next="2"/>
              <script id="2" next="3" >
                <![CDATA[
                  finishCall = true;
                ]]>
              </script>
              <subdialog id="3" next="4" name="goodbyStatement"/>
            <end id="4"/>
          </catch>
    <end id="7"/>
</subdef>
```

Figure 8 - Procedural code in a script block

Future enhancements

Inclusion of spoken interfaces optimization techniques or *best practices* into custom-developed systems is not straightforward. A big advantage in using an interpretable and high-level dialogue representation language like the one proposed in this work is that such "dialogue practices" can be incorporated seamlessly into the underlying dia-

2 VoxBuilder. http://www.voxbuilder.com/

logue interpretation logic, removing the burden from the dialogue developer.

Furthermore, a complete project management support is foreseen, where a project involves a dialogue and its composing subdialogues, together with definition of templates. High level templates serve as guidelines in the development of abstract tasks, such as assessing the patient's psychological stage.

Currently, we have a strong commitment on the integration of a more elaborated semantic interpretation by integrating AdaRTE with a NLP application that supports lexicalized grammars to increase expressivity. In this way, not only recognition does not depend on the grammars supported by VBs, but also more natural dialogues will be supported improving the patient's perception of the dialogues.

Other components such as integration with Workflow Management Systems and facial expressions and gestures realizers must also be considered in future research.

Conclusion

We have presented an architecture for next-generation dialogue interpretations and successfully built an engine for easily dialogue deployment. AdaRTE supports the generation of the standard VoiceXML to communicate to VBs. As real-world test cases, we have reengineered two healthcare dialogue prototypes by using the novel architecture and showed that dialogue development and deployment times are remarkably optimized with respect to customized coding.

The AdaRTE system is foreseen not only as a reliable platform for dialogue deployment, but also as a framework for incorporating advanced features of speech recognizers, including increased support to adaptability, and natural language understanding and generation.

Acknowledgments

The authors thank Dr Robert Friedman (Medical Information Systems Unit, Boston, MA) for information on the TLC-COPD dialogues, and Timothy Bickmore (College of Computer and Information Science, Northeastern University, Boston, MA) for the discussions that led to the AdaRTE project.

References

[1] Migneault JP, Farzanfar R, Wright JA, Friedman RH. How to write health dialog for a talking computer. Journal of biomedical informatics. 2006 Oct;39(5):468-81.

[2] Piette JD. Interactive voice response systems in the diagnosis and management of chronic disease. Am J Manag Care. 2000;6:817-27.

[3] Corkrey R, Parkinson L. Interactive voice response: review of studies 1989-2000. *Behav Res Methods Instrum Comput*. 2002;34:342-53.

[4] Krishna S, Balas EA, Boren S, Maglaveras N. Patient acceptance of educational voice messages: a review of controlled clinical studies. *Methods Inf Med*. 2002;41:360-9.

[5] Young M, Sparrow D, Gottlieb D, Selim A, Friedman RH. A telephone-linked computer system for COPD care. Chest. 2001;119:1565-75.

[6] Giorgino T, Azzini I, Rognoni C, Quaglini S, Stefanelli M, Gretter R, et al. Automated spoken dialogue system for hypertensive patient home management. International Journal of Medical Informatics Elsevier. 2004;74:159-67.

[7] Bickmore T, Giorgino T, Picard R. Guest Editorial: Special Issue on Dialog Systems for Health Communication. Journal of Biomedical Informatics Elsevier. 2006;39:465-7.

[8] McGlashan S, Burnett DC, Carter J, Danielsen P, Ferrans J, Hunt A LB, Porter B, Rehor K, Tryphonas S. Voice Extensible Markup Language (VoiceXML) Version 2.0. 2004 [cited; Available from: http://www.w3.org/TR/voicexml20/

[9] Mittendorfer M, Niklfeld G, Winiwarter W. Evaluation of Intelligent Component Technologies for VoiceXML Applications; 2001.

[10] Black LA, McTear M, Black N, Harper R, Lemon M. Di@log by design: An Integrated Tele-Care Communication Infrastructure. In: Proceedings of the 26th Annual International Conference of the IEEE EMBS San Francisco 2004;2:3290-3.

[11] Mittendorfer M, Niklfeld G, Winiwarter W. Making the voice web smarter-integrating intelligent component technologies and VoiceXML. In: Proceedings of the Second International Conference on Web Information Systems Engineering 2002;2:126-31.

[12] Hataoka N, Obuchi Y, Mitamura T, Nyberg E. Robust speech dialog interface for car telematics service. In: First IEEE Consumer Communications and Networking Conference 2004:331 - 5.

[13] Hartman J, Vila JA. VoiceXML Builder: A workbench for investigation voice-based applications. In: Proceedings of the 31st Annual Frontiers in Education Conference, 2001 2001;3:S2C 6-9.

[14] Balentine B, Morgan DP. How to Build a Speech Recognition Application. A Style Guide for Telephony Dialogues. Second ed: EIG Press 2001.

[15] McGlashan S, Hunt A. Speech Recognition Grammar Specification Version 1.0. 2004 [cited; Available from: http://www.w3.org/TR/speech-grammar/

Address for correspondence

E-mail address: linamaria.rojas@unipv.it
(L. M. Rojas-Barahona).

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Multi-channel Physiological Sensing of Human Emotion: Insights into Emotion-Aware Computing using Affective Protocols, Avatars and Emotion Specifications

Panagiotis D. Bamidis, Andrej Luneski, Ana Vivas, Christos Papadelis, Nicos Maglaveras, Costas Pappas

Lab of Medical Informatics, The Medical School, Aristotle University of Thessaloniki, Greece

Abstract

This paper introduces a methodology for combining multi-channel psycho-physiological recordings of affective paradigms into a framework where the scientific results of such experiments are utilized in the human computer interaction context to model the computer's response based on the emotional context of the user and the situation. An affective protocol is described the results of which are expected to be combined with anthropomorphic avatars that enhance the man-machine interaction. The technological infrastructure of the later component is provided by means of XML specifications of signal descriptions and emotion recognition, as well as avatar behavior generator descriptions.

Keywords:

affective computing, emotion identification, specification, avatar technology.

Introduction

The interaction between humans and computers (HCI) has been a subject of research and discussion for a long time. The goal in the enhancement of HCI is getting it closer to the interaction between humans [1]. The essential human ability, and the one that nearly all the research is focused on, is human intelligence and its incorporation into computers. The HCI has experienced the introduction of facial and speech recognition, natural language recognition, as well as, software intelligent agents than can learn and reason on their actions [2]. However, until recently, emotions were rarely the topic mentioned in the discussions of human intelligence and their application in computers was not even considered. The reason was that emotions were considered as extraordinary human ability. Nowadays, however, there are many proofs of the importance of the emotions in the expression of intelligence [3]. The arguments put alongside the significance of emotions gave birth to a new area of "emotional intelligence", defined as "the capacity to understand emotional information and to reason with emotions" [3]. All the work that is done with computing and is related to, arises from or deliberately influences emotions on computers is called "affective computing" [4]. Affective computing has been a hot topic in recent years mostly because it introduced a new area in computing and therefore an increasing number of applications in that area. Until recently research has been mostly focused on monitoring user emotional reactions and trying to distinguish between different emotion categories such as fear, anger, sadness, happiness etc. The ability of recognizing human emotions requires the computer to monitor the user and based on certain parameters or conditions classify his/her emotional state. Since we are intending to copy the essence of human-human interaction, it is mandatory to identify the natural way of accomplishing this task by humans. Namely, human beings use several ways of communication and emotion recognition using their biological sensors of sight, touch and sound. Facial communication has proven to be of unique importance since facial expressions help in a great deal in identifying people's emotions [5]. Moreover, vocal communication as well as gesture based can be successful, to certain extend, in sensing and distinguishing between several classes of emotions.

Apart from using human natural senses, emotions can be recognized by monitoring psycho-physiological changes in the user [6, 7]. Naturally, for this to take place certain sensors have to be used recording heart pulse/activity (e.g ECG), galvanic skin response (GSR or SC), respiration rate, electroencephalogram (EEG) etc. This type of emotional reaction recording has been most attractive recently and it is the main focus in this project as well. While in the field of medicine the influence of emotions on the human health is of major concern, in human-computer interaction the focus is on understanding the user's emotions and improving the quality of software programs accordingly.

Understanding emotions is far from simple due to the complexity of human physiology. It would be so if each emotion category was characterized by a specific physiological pattern. However, certain emotions, even as different as love and fear, can cause similar effects on human physiology and, therefore, may be misrecognized. The emotion recognition process follows a sequence of steps even for the preparation of data before the recognition process. The initial step is to acquire the data from the user physiological signals. Following is the extraction of the key features of the signals [8]. The features form the basis of the comparison method. The feature extraction procedure can vary and depends on the goals of each specific project and therefore cannot be discussed in general. The final step in emotion recognition is performing

the classification of the emotional data into emotion categories based on specific classification techniques [9]. Mostly used pattern classification methods are the Hidden Markov Model [10], Fisher linear Projection [11], Support Vector Machines [9] etc. However, these techniques are appropriate for classification of large amounts of data into categories. Moreover, there is a clear need for collections of representative emotional signals in short samples of 20-50 seconds containing digital signal data, in a flat file format such as that used in the MIT-BIH file library [12]. There exists also a need for a more representative way of specifying an emotion into one data record. An XML file containing the required signal data was successfully used in [9], thereby providing evidence that an XML based representation of the emotion elements is a suitable format of data specification. Furthermore, the contents of the XML data record are of great importance for interconnection among different research results, platform independency and reusability purposes, which can be used in telemedicine, decision-making etc. [13]. Thus, it is essential to follow certain standards or widely accepted guidelines for the structure and names of the elements in the record. One related standard was introduced in [14]. Basically, it is an introduction of a markup language – ecgML for modeling and storing ECG data of patients based on XML representations.

Significant progress has been made in all the above fields and therefore new ideas arose which basically provide answers on the question "What can the computer do after recognising human emotions?". Lisetti and colleagues [15] have recently investigated the use of agent-centered modalities or modes (avatars), and multimodal feedback given to a system user. For example, an interface agent for an e-health system session can display empathy via an anthropomorphic avatar who adjusts its facial expressions and vocal intonation according to the user's emotional state, as the latter is depicted by the set of measurements.

In the light of the previous developments, the scope of this paper is twofold. First, to present a step-wise approach to the design of an experimental protocol that aims to enable multi-channel physiological sensing of a subject's emotion. The second goal is to prepare the theoretical and technological grounds for later direct adaptation of computer user interfaces guided by the elicitation and identification of emotions. The former goal is attained through the use of physiological sensors like EEG, ECG, and skin conductance, while the latter one is mainly driven by the notion of avatar technology and the preparation of the relevant technological platform that will exploit the results of the experimental analysis.

Material and methods

The AFFECTION project context

This piece of work is part of a collaborative project, called AFFECTION, between the Lab of Medical Informatics at the Medical School of the Aristotle University of Thessaloniki, Greece, and the Brain Science Institute of RIKEN in Japan. The project aims at creating a scientific foundation for the robust identification of human emotional states through fusion and correlation of data from a multi-sensor research framework. The emotion-related research findings will be subsequently incorporated within usability evaluation methodological frameworks to enable new objective/direct evaluation methodologies, as well as, new interface adaptation strategies. The project is envisaged to contribute to dynamic characterization and recognition of the subjects' emotional state upon interaction with computer systems. To achieve this, the project will employ multimodal recordings such as vocal expressions and physiological signals of the autonomic (heart rate, blood pressure, skin conductance etc) and central nervous system (EEG, MEG). Research will be carried along the following main stages: in the first, procedures, examples and baselines of multiple-channel emotional responding approaches are utilised to obtain fundamental inter-correlations of the various physiological measures in the light of behavioral data. The last stage involves the use of specific paradigms in order to incorporate the findings into a unified framework that ascertains the changing emotional state of the user and allows for adaptation of computer user interfaces based on the user and application context. A block diagram of the project is given in Figure 1.

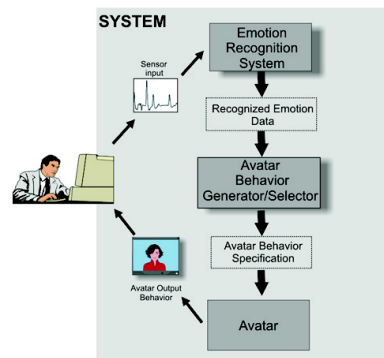


Figure 1 - A block diagram of the AFFECTION system

Affective protocol design

As already mentioned, emotions are complex phenomena, which include a wide range of observable behaviours, expressed feelings that are private and subjective, and changes in body states (distinctive somatic and autonomic responses). To achieve a first insight and start building on the scientific foundation of the AFFECTION project, the initial study aims at investigating the relationships between the different patterns of brain activation, autonomic responses, subjective experience and cognitive behavior related to emotionally-evocative photographs from the International Affective Picture System (IAPS) [16]. 50 healthy adults ranging in age from 18 to 40 years old form the participants' sample. Before the experiment, participants are updated and asked to sign a consent form. A full self evaluation questionnaire is given to them right afterwards, in order to have a (subjective) definition of the emotional state in general.

In Phase 1, we study the psychological and neurophysiological correlates of passive exposure to emotionally-

evocative stimuli. Four blocks of emotionally-evocative stimuli are selected from the IAPS, taking into account mainly two major dimensions, pleasure and arousal. Pictures are selected from the "Affective Space" as defined by the mean rating on the Valence (pleasure) and Arousal dimension. Each condition of "Affective Space" is manipulated between blocks, and the order of the Affective Space-block is counterbalanced across participants to avoid the order effect. In each trial, each stimulus is presented during 1 second. No inter-stimulus interval is scheduled for these blocks. There are 40 trials for each block, randomly selected from a larger set of pictures from each affective space condition. After each block, subjects fill in a questionnaire indicating how they feel right after this sequence block. In the second phase, the same pictures are used in a visuospatial attention paradigm, in order to investigate the effect of emotional processing on cognitive behavior. The independent variables are Affective Space block condition, and the visual field (left or right). On each trial, a central fixation cross appears for 500 ms followed by two pictures (one with high or low mean rating on valence and arousal and the other one with middle mean rating on both dimensions) for 500 ms. Then a target (a small asterisk) appears on the location of the emotional picture or in the location of the neutral picture (see Figure 2). The target remains until a response is made or until 2000 ms elapse. Participants are required to detect the appearance of the target; this can be done by pressing a key. During the experiment, recordings of the 10-20 EEG are conducted. In addition, simultaneous recordings of GSR, EOG, ECG are also taken.

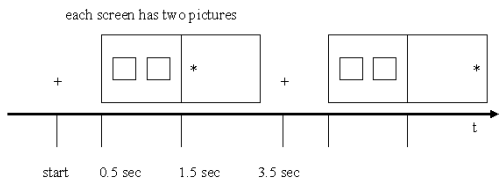


Figure 2 - Phase 2; block sequence of the affective protocol

The emotion recognition subsystem

Figure 3 shows a Logical View of the system Architecture.

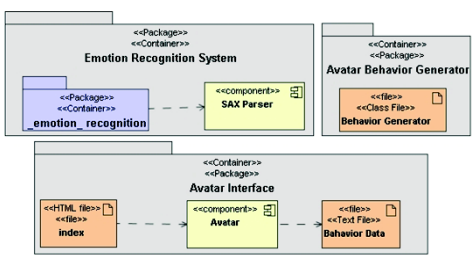


Figure 3 - Design model – logical view of the architecture

The emotion recognition system component is the most important element of the system since effective task execution of all other system components depends on it. It is the first component of the execution sequence and the component that interacts with the environment outside the system i.e. the component that accepts the input from the user physiological signals. As mentioned before, that input comes in the form of XML data specifications. There are two input requirements that this system needs. The first is the filename of the XML representation for the input emotion to be recognized and the other is actually the invocation of the recognition procedure. The operation of this subsystem can be divided into three main tasks, namely, reading user data; recognizing the emotion by comparison to its knowledge base; outputting the recognizing emotion into a proper format.

The emotion recognition system currently requires different types of physiological data in order to perform a successful recognition on the user emotion: Electromyogram (EMG), Blood volume pressure (BVP), Skin conductivity (SC), Respiration rate (RR), Electrocardiogram (ECG). It is envisaged to include EEG in the near future to fully accommodate the needs of the affective protocol analysed above.

Apart from the signal data on which the actual recognition method is applied, additional information is required by the system. This is: (i) User identification (e.g. personal information, possible medical information etc), (ii) Signal sample data (e.g. time length, measuring unit etc.), and (iii) Measurement information (recording time, date, place, temperature etc).

Since the input to the emotion recognition system is an XML file having all the above required data, an internal part of this system is a parsing functionality that extracts only the data needed for the recognition process.

Emotion recognition is the key feature of this system and the "reasoning" part of it. The reasoning is actually a comparison between the emotion data obtained from the user and the emotions that the system can recognize. The number of these emotions can be variable and depends merely on how many XML data files are stored, each representing one "classical" emotion such as fear, happiness, sadness etc (Figure 4). The comparison is done between each of those emotions and the input emotion data by comparing the four signals independently. A simple method of comparison is chosen at the time being, as the scope is to merely test the feasibility of the overall methodological approach. To be more precise, following an extensive literature review on emotion recognition using pattern recognition techniques, sample signals were extracted for several emotions (anger, fear etc.). These samples can be considered as characteristic signals of the appropriate emotions and further matching can be done by simply comparing any new signal to each of them. The comparison is done using a weighted version of the dot product (cosine matching) technique on data samples of the signal. For each signal type we chose a number of characteristic features like the mean value, the number of peaks, Average amplitude of the peaks etc.

Avatar behavior generator

This subsystem is quite simple. Its one and only goal is to generate a file with specific parameters that will instruct the avatar behavior. The reason for considering it as a separate component is its functionality. The reaction of the entire system to the input emotion depends on this component, and for experimentation purposes in this project it is created with a simple classification function. The basic idea of the classification function is to discover the level of emotional reaction and to connect that to the output behavior. For example, if the emotion recognition system gave as output that it recognized anger at user side and with 80% correctness or the user was feeling very angry, this component needs to find out which parameters are the appropriate for such a strong emotional reaction on the user side. This component takes two inputs from the emotion recognition system: the recognized emotion and the comparison value. There is much freedom in the decision for the output behavior. In different studies or experiments, the implementation of this component varies. In this first project demonstration, it was decided only to mimic the emotional reactions of the user, in the facial area, since the avatar represents only facial human-like characters. The output is a file that has a predefined structure, since the avatar is deployed into this system and has specific requirements that cannot be changed. The Haptek platform was used for the avatar creation. The avatar has to be embedded into a web page and thus there are two kinds of inputs. First, is the input that the user gives to the web page for running the avatar; the second input is the file containing the behavior parameters. The web page contains the avatar and one button for starting the avatar behavior.

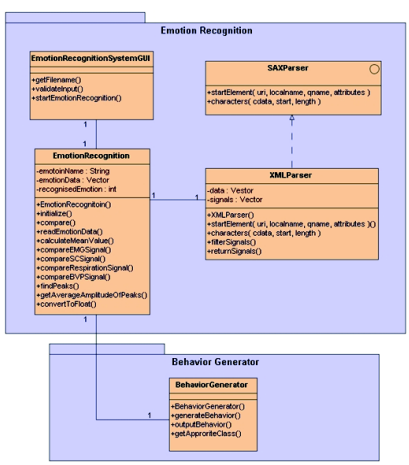


Figure 4 - Design model (class diagram) – emotion recognition container and behavior generator

Results

The emotions stored in the emotion recognition system, as well as, the input emotions are represented as XML files. The file is a modified representation of the standard for ECG data representation described in the literature review. The main element in the file is the EmotionRecord. It contains four records, each corresponding to the physiological signals, taken from the user emotion for example, EMG, BVP, SC, and RR in the current version. Furthermore, the record element consists of the record data (the main element) and some additional information such as the recording device, the recording date etc. Figure 5 shows a portion of the XML specification. An example of the comparison based on the features of the signals is given in Figure 6 only for the skin conductance case, since the implementation of the rest of the features is quite similar. The avatar implementation is a combination between the HTML page representing the interface with the user, the scripting files with the avatar code and the external data file for the avatar behavior. The appearance of the avatar is possible by installing the Haptek Player plug-in, which can run the Haptek HyperText commands as the interface between the HTML and the JavaScript functions. The behavior of the avatar and its output to the user is triggered by pressing a kind of "Run button" at the moment, but it is envisaged to be automatically triggered in future versions.

```
…
<EmotionRecord studyID="ECG00001">

      <Record type="EMG">
      <!--Recording device can be here-->
      <RecordData>
            <Channel>MLII</Channel>
            <Waveforms>
            <XValues>
            <XOffset dataType="time">00:00:00.000</XOffset>
            <Duration dataType="time">00:05:06.000</Duration>
            <SampleRate unit="Hz">20</SampleRate>
            </XValues>
            <YValues>
            <RealValue>
            <From dataType="time">00:00:00.000</From>
            <To dataType="time">00:00:10.000</To>
            <Data>1.15,2.50,2.10,2.15,2.10,2.15 … </Data>
            </RealValue>
                  </YValues>
            </Waveforms>
      </RecordData>
      </Record>
```

Figure 5 - XML specification of one record

```
public void compareSCSignal()
{
      Vector storedSCSignal = (Vector)emotion.elementAt(2);
      float x=0;
      int numberOfSCR = 0;
      int numberOfPeaks = 0;

      for(int i=1; i<(storedSCSignal.size()-1); i++)
      {
            if( (Float)storedSCSignal.elementAt(i) >
         (Float)storedSCSignal.elementAt(i-1) &&
               (Float)storedSCSignal.elementAt(i) >
         (Float)storedSCSignal.elementAt(i+1))
            {
                  numberOfPeaks+=1;
                  if(checkForSCR(storedSCSignal,i) == true)
                  {
                        numberOfSCR+=1;
                  }
            }
      }
}
```

Figure 6 - The compareSCSignal method for SC features

Discussion

The intention of this paper was neither to introduce a new way of identifying an emotional signal, nor to put across a new affective protocol, but rather to carefully explain the individual methodological steps for the scientific exploitation of affective computing solutions. To this end, the

affective protocol design, currently under data collection, will provide the means to build up a knowledge base of emotional signals. One of the main strengths of this project was the idea for XML specifications of emotional data for the user. This can introduce new ways of emotion recognition by qualitative classification of emotions and not only by statistical quantitative experimentation. Furthermore, the contents of the XML data representation are represented by a modified standard for ECG patient data representation used elsewhere.

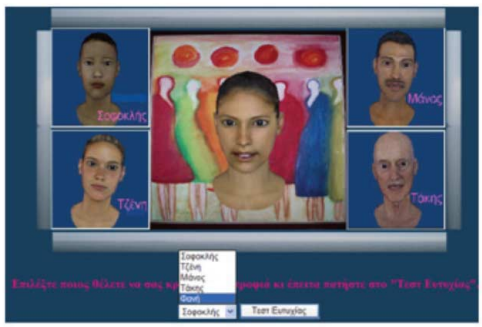


Figure 7 - Avatar appearance after emotion identification

The standard representation was modified to satisfy the needs of the emotion XML specification for this project. Obviously, the current XML specification of emotion data can be significantly improved by adding elements that will more accurately describe physiological signal data rather than only having numerical signal representations (e.g. EEG feature descriptions). Furthermore, the emotion recognition method based on the dot product comparison of the features can become more consistent so that it represents a proven classification of the emotions based on the results of the psychophysiological (affective) experiments. The generator of the avatar behavior is the component that can be also easily extended. Improvements can be envisaged in the "reasoning" method for generation of behavioral parameters. In specific, if the component is used in relation to larger amounts of emotion categories, it will be obviously more efficient and it can produce more classes of behavior parameters appropriate for the specific emotions. Finally, the avatar itself might go through several enhancements in its appearance to the user. The ultimate goal of embedding emotional awareness into the computer is to produce a system that can recognize emotions, and respond intelligently and appropriately in real-time, just like humans do. This paper provides a methodology for enabling the construction of more intelligent systems based on scientific reasoning and experimental results and not mere technological artifacts. The envisaged incorporation of the findings into a unified framework that ascertains the changing emotional state of the user and allows for adaptation of computer user interfaces based on the user and application context seems an exciting future prospect of this project. The wide variety of application areas that can be associated with such a system e.g. interactive games, learning systems, e-health/home care systems, etc call for a careful continuation of the project development.

Acknowledgments

This work has been benefited from a grant from the Greek GENERAL SECRETARIAT FOR RESEARCH & TECHNOLOGY.

References

[1] Reeves B, Nass C. The Media Equation: How People Treat Computers, Cambridge Univ. Press, 1996.
[2] Ghauoi C. Human computer interaction encyclopedia, Idea Group Inc., London, UK, 2006.
[3] Mayer JD, DiPaolo M, Salovey P. Perceiving affective content in ambiguous visual stimuli: a component of emotional intelligence. J Pers Assess 1990 ;54(3-4):772-81.
[4] Picard RW. Affective Computing. MIT Press, Cambridge, USA, 1997.
[5] Ekman P. Emotion in the Human Face, Cambridge University Press, New York, 2nd edition, 1982.
[6] Hone K. Empathic agents to reduce user frustration: the effects of varying agent characteristics. Interacting with Computers 2006; 18(2):227-45.
[7] Bamidis PD, Papadelis C, Kourtidou-Papadeli C, Pappas C, Vivas AB. Affective computing in the era of contemporary neurophysiology and health informatics. Interacting with Computers 2004; 16(4):715-21.
[8] Kim K, Bang S, Kim S. Emotion Recognition System Using Short-term Monitoring of Physiological Signals. Med. Biol. Eng. Comp. 2004; 42:419-27.
[9] Katsis CD, Ganiatsas G, Katertsidis N, Fotiadis DI. Emotional Understanding Using A Wearable Multisensorial Platform. In Bamidis PD et al (Eds), Improving the Quality of Health Information: an international perspective, Proc. iSHIMR 2005:188-97.
[10] Scheirer J, Fernandez R, Klein J, Picard RW. Frustrating the user on purpose: a step toward building an affective computer. Interacting with Computers 2002; 14(2):93-118.
[11] Healey J and Picard RW. Digital Processing of Affective Signals; In Proceedings of ICASSP, Seattle, 1998.
[12] MIT PhysioNet. PhysioNet: the research resource for complex physiologic signals. http://www.physionet.org/
[13] Varni A, Kemp B, Penzel T and Schlogl A. Standards for biomedical signal databases. IEEE Eng Med Biol Mag. 2001;20(3):33-7.
[14] Wang H, Azuaje F, Jung B, Black N. A markup language for electrocardiogram data acquisition and analysis (ecgML). BMC Medical Informatics and Decision Support 2003; 3(4).
[15] Lisetti C, Nasoz F, LeRouge C, Ozyer O, Alvarez K. Developing multimodal intelligent affective interfaces for tele-home health care. Int J Hum-Comput Stud 2003; 59:245-55.
[16] Bradley MM, Cuthbert BN, Lang PJ. Picture media and emotion: effects of a sustained affective context. Psychophysiology 1996; 33(6):662-70.

Address for correspondence

Panagiotis D. Bamidis,
Lab of Medical Informatics,
Medical School,
Aristotle University of Thessaloniki,
P.O. Box 323, 54124,
Thessaloniki, Greece.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

A Framework for Cognitive Monitoring Using Computer Game Interactions

Holly B. Jimison, Ph.D.[a,b], Misha Pavel, Ph.D. [a,b], Payton Bissell, B.S. [b], James McKanna, B.S. [b]

[a] *Oregon Health & Science University, Portland, Oregon, USA*
[b] *Spry Learning Company, Portland, Oregon, USA*

Abstract

Many countries are faced with a rapidly increasing economic and social challenge of caring for their elderly population. Cognitive issues are at the forefront of the list of concerns. People over the age of 75 are at risk for medically related cognitive decline and confusion, and the early detection of cognitive problems would allow for more effective clinical intervention. However, standard cognitive assessments are not diagnostically sensitive and are performed infrequently. To address these issues, we have developed a set of adaptive computer games to monitor cognitive performance in a home environment. Assessment algorithms for various aspects of cognition are embedded in the games. The monitoring of these metrics allows us to detect within subject trends over time, providing a method for the early detection of cognitive decline. In addition, the real-time information on cognitive state is used to adapt the user interface to the needs of the individual user. In this paper we describe the software architecture and methodology for monitoring cognitive performance using data from natural computer interactions in a home setting.

Keywords:

Computer monitoring, cognitive assessment

Introduction

Cognitive performance is a key health concern of elders in the United States. In fact, maintaining cognitive health is often the most important factor in being able to age in place. Nearly 50% of all people over the age of 85 are found to have a measurable decline in cognitive function [1]. However, common clinical practice does not offer methods for detecting cognitive decline at an early stage, when therapies may be more effective. Recent research has demonstrated the importance of detecting cognitive decline in an early stage [2-4]. Some cognitive issues have immediately treatable causes, such cognitive disturbances due to medication interactions or short-term medical conditions. However, even with long-term conditions, such as dementia, there are many new therapies that researchers presume would have improved efficacy with earlier detection. In this paper we describe a framework for using unobtrusive computer interaction data to infer cognitive changes on the part of computer users. Frequent assessments allow us to detect relevant changes in various aspects of performance that can be used to adapt the user interface in real time and also provide a mechanism of early detection of cognitive problems.

Growing use of computers by elders

Elders are the fastest growing demographic of new computer users in the United States. In a recent survey conducted by the Pew Internet and American Life Project [5], they found that 22% of American adults over the age of 65 use the Internet. Interestingly, elders in this group are even more likely than other Internet users to go online and check email each day [5]. In addition, nearly 35% of elders who use a computer have played a game online, comparable to 39%, the average rate of computer game play for other age groups. Given this rapid growth of computer use by users at risk for cognitive problems, as well as the current large use of computers by the advancing wave of baby boomers, we have an important opportunity to collect and interpret naturalistic computer interaction data for diagnostic purposes. In this project on cognitive monitoring using computer interaction data, we have focused on the interpretation of interactions in computer games that we have specifically designed to probe cognitive performance.

Current methods of cognitive assessment

In standard clinical practice, cognitive screenings are usually performed only at advanced age or if there are already patient or family concerns about cognitive dysfunction. These screening tests, such as the Mini-Mental State Exam, the Kokmen Short Test of Mental Status, and the Memory Impairment Screen, can be performed in a physician's office, but are fairly coarse and not particularly useful for the early detection of problems [6]. More complete neuropsychological batteries can be performed to obtain more sensitive diagnostic information. These normally include measures of short-term and working memory, divided attention, motor speed, planning, and general executive function. In our project on designing computer games for cognitive monitoring, we attempted to incorporate proxies for the following standard tests of cognitive performance:

1. **Verbal Fluency -**The test is focused on semantic processing and recall from long term memory (LTM). The test procedure requires the participants to recall as many words as possible given a specific semantic category or one or more phonemic constraints.

Selected for best paper award.

2. **Word-List Acquisition and Delayed Recall -** The test is focused on learning and recall from short term memory (STM) as well as LTM. The test procedure requires the participants to learn and recall a list of words with three trials and then after an intervening task.
3. **Word list Recognition -** This is a test of the ability to recognize words previously presented during the Word-List Acquisition test. The participant is asked to discriminate between the words that were presented in the list from distractor words. Together with the Word-List Acquisition test, the recognition test can distinguish whether the "forgotten" items were truly lost or the memory trace was just too weak to support reliable recall.
4. **Trail-Making Test** - This test is focused on complex visual scanning, mental tracking and mental flexibility. The participants are asked to trace a sequence of digits and then a sequence of interposed digits and letters.
5. **Symbol Digit Modalities Test -** This test is used to assess the ability to sustain attention and to perform coding task. The participant is given a table associating a simple symbol with each digit and then is asked to assign a number to each of a long list of these symbols.
6. **Digit Span -** The focus of this test is working memory and sustained attention. The participant is asked to repeat a sequence of digits, starting with short sequences and then of increasing length. The following task is to do the same thing in reverse order.
7. **Finger tap test -** Although this test is focused on the speed of motor control, there is increasing evidence in the literature that this type of test is useful in predicting future decline in cognitive abilities. The participant in this test is asked to push a switch as many times as he can within a ten second interval. One feature of this test is that the results of the performance are insensitive to educational level and other demographic variables.

These neuropsychological tests are usually performed by trained psychologists and usually done no more frequently than once per year. One of the hallmarks of cognitive impairment is the increasing variability in performance. Infrequent assessments do not offer a mechanism to pick this up. In fact, the results of standard cognitive measures are clouded by a need to reference the performance metrics directly to population norms. Many cognitive tests are highly affected by differences in educational level, language abilities, etc.

In our work with monitoring computer game interactions to infer cognitive performance, we make use of these metrics of verbal fluency, short-term and working memory, planning abilities, and divided attention. However, we are able to make assessments every time an elder uses a computer. Although our computer assessments are less direct and less controlled than the standard tests, we have the benefit of multiple nearly continuous measures and can analyze within-subject trends. This substantially reduces unwanted confounding effects due to education, language abilities, and culture. In addition, we are able to characterize variability in performance over time, which in itself is a powerful indicator of cognitive function.

Materials and methods

Unobtrusive monitoring of computer game interactions

In our project on monitoring elders' computer interactions, we first performed a needs assessment to define elders' preferences for computer applications, games, and potential barriers to computer use. We used focus groups and surveys to help us define a set of features for an elder Web portal that we could use as a research environment to collect real-time interaction data. We also defined a set of enjoyable computer games that could be adapted for cognitive monitoring. To select the games for further development, we observed which features were most enjoyable and easily understood by elders and then also did a cognitive task analysis on each of the games to characterize its appropriateness for providing information on one of the cognitive dimensions described in the previous section on standard cognitive tests.

We currently monitor all keyboard and mouse interactions, both within game play, and in conventional computer applications. Each of the adaptive cognitive computer games for elders are designed to measure various aspects of the standard cognitive tests described in the previous section. For example, two of our computer games are designed to measure verbal fluency (e.g., ability to name as many animals as possible within 60 seconds).

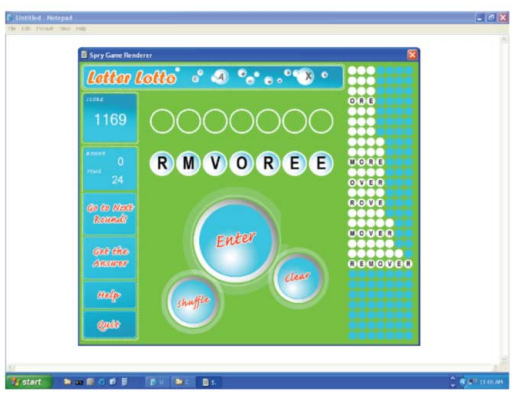


Figure 1: A word jumble game where we measure the user's relative ability to find longer and more complex words from a set of 7 letters.

Figure 1 shows an example of a word jumble game, where the users are give a set of 7 letters and asked to generate as many words from that set as quickly as they can. They are given cues on the right of the screen to show how many words are possible. Entering a word using all 7 letters allows them to go onto the next round. Two basic metrics relating to verbal fluency can be generated from monitoring the users' interactions in this game: 1) speed of word generation and 2) the complexity of the words generated (defined by word length and frequency of use in the English language).

Figure 2: A word game where we measure the user's relative ability to find longer and more complex words in a difficult search environment.

Figure 2 shows another word game designed to measure verbal fluency, but with the additional task of search and planning. In this game, the user must connect adjacent letters to form words as quickly as possible. A higher score is given for longer words and for using highlighted letters. With the reference abilities of verbal fluency measured in the game shown in Figure 1, we are now able to quantify the increased performance requirements due to search and planning. In each of these games, the difficulty of the board layout is adapted to the skill of the user. Our adaptation algorithms keep the success rate at approximately 60%-80%, so that users are challenged, but not frustrated. This also gives us the best opportunity to measure performance, i.e., users' scores do not "top out" when the task is too easy or "bottom out" when the task is too difficult.

In these games, the user's creation of longer and more sophisticated words (against time and difficulty of available letters) we rate as having higher verbal fluency. We concentrate on monitoring relative performance (with respect to the user's baseline) to look for differences. This is likely to be a more sensitive measure that is less influenced by education and language abilities, and more influenced by cognitive changes.

For a direct measure of short-term and working memory, we adapted the Concentration card game to the computer, as shown in Figure 3. Users must remember the location of various cards they select and match pairs. Game difficulty is adapted based on number of cards and the cognitive difficulty of the matches. These range from simple shape and color matches to cognitively more difficult matches, such as matching a digital clock time with the analogue picture equivalent. For this game, we estimate memory ability using an adaptive memory buffer metric. This metric is defined by filtered estimate (similar to a moving average) of the maximum number of card flips back that a user can successfully remember seeing a target card.

We have designed other computer games to specifically test the remaining dimensions of cognition. Figure 4 shows a shape and color matching game that provides us with measures of planning (inferring the number of steps ahead a user would have to be able to plan in order to be successful). In this game we can also manipulate difficulty and provide added features to test memory and divided attention.

Figure 3: Example of a memory computer game.

Figure 4: Color and shape matching game that tests planning ability, memory and attention.

In addition to the games where we adapted activities that elders already found to be enjoyable, we also took standard tests like the Trail Making Test and adapted it to be fun. In this game, users use the mouse input device to select circles in numerical sequence, letter sequence, and mixed targets. This activity is similar to the Trail Making Test in requiring several dimensions of cognitive executive function, including visual search, attention and set switching. Each of these additional components can be extracted and assessed with repeated game use. These more frequent measures allow us to monitor within-subject trends and also monitor performance variability. Both of these features offer promise in being able to detect cognitive problems earlier, and potentially, more reliably.

Results

Most of our experience and testing of computer games for cognitive monitoring has come from our work with an implementation of the popular Solitaire game of FreeCell, as shown in Figure 5. We found that this game was by far the favorite with the elders that we interviewed and it was the first computer game we adapted for use in cognitive monitoring. In our research version, we compare user performance to our computer solver. The lower graph of Figure 5 shows the game difficulty starting at 82 moves to

optimal solution, with the lower line showing the computer solver's direct path to solution. The upper line shows the subject's moves going toward and away from best solution. We use the slope of the subjects performance as a measure of efficiency of play. In our early pilot work comparing FreeCell performance of cognitively healthy elders to those with diagnosed mild cognitive impairment, we were able to use the efficiency metric to distinguish the two groups.

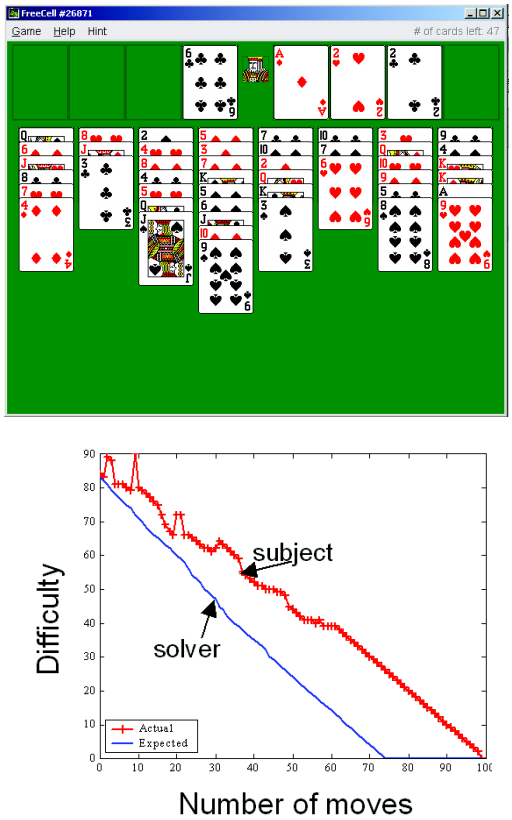


Figure 5: Sample game of FreeCell (Solitaire game requiring significant planning) and a diagram showing subject performance vs computer solver.

Table 1 shows the results of our early pilot tests to show the feasibility of monitoring computer interactions in the home. We monitored 12 elders in a local senior residential facility for a period of 3 weeks. Using conventional neuropsychometric tests described earlier, we found that 3 of the elderly subjects (mean age 80.2 +/- 8.0) had mild cognitive impairment. Using only data from their FreeCell performance we were able to distinguish cognitively healthy subjects from those with mild cognitive impairment. Interestingly, the variability of the measures over time was in itself a useful feature in classifying cognitive impairment.

Table 1: FreeCell efficiency cognitive metric scores for 9 cognitively health elders and 3 elders with mild cognitive impairment.

	Ave of Subjects' Ave Efficiency	SD of Subjects' Ave Efficiency	Average of Subjects' SD Efficiency
Normals	0.58	0.12	0.38
MCI	0.27	0.72	0.55

Subsequent to the pilot test of the FreeCell game, we deployed the full set of 9 adaptive cognitive computer games into the homes of 30 elders. The average age of our participants is 80.4 ± 6.0 years. Most are female (83%) and have an average level of 15.2 ± 2.7 years of education. In this evaluation, we have demonstrated that we are able to extract cognitive measures from routine game play on the part of elders. The play the games successfully and on debriefing, report that they enjoy the experience. The plot in Figure 6 shows their game usage patterns over a period of 3 months.

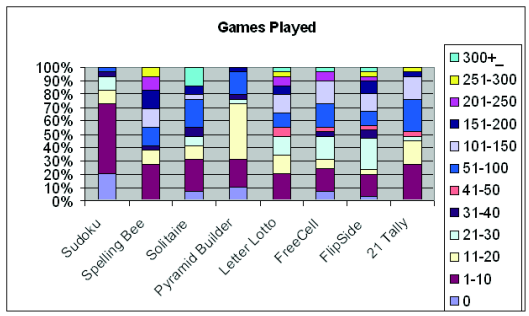


Figure 6: Computer game usage by 30 elders over a period of 3 months.

Software Architecture for Cognitive Monitoring

We have developed a rich set of tools for assessing cognitive performance based on the unobtrusive collection of computer interaction data. Our measures are based on keyboard and mouse interactions for both cognitive computer games and conventional applications. The measures include metrics of verbal fluency (word processing and word games), motor speed (login typing, game speed), memory, attention, planning and general executive function. Figure 7 shows our general software architecture for collecting and analyzing the monitoring data.

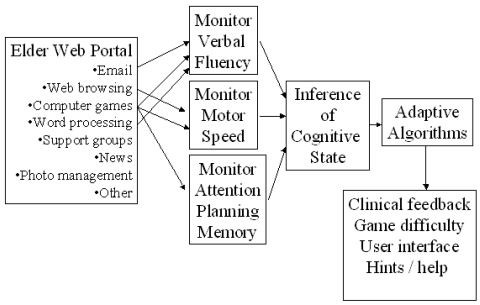


Figure 7: Overview of software architecture for cognitive monitoring.

Real-time analysis of game data takes place on the elder's local machine. This data is used to adapt the difficult of the ongoing computer game (in some cases where appropriate) and also used to adapt the level of difficulty for a user's upcoming games. We also use real-time analysis and feedback to tailor hints and help messages as part of the user interface. If we realize that a user is having memory problems or divided attention problems, we are then able to immediately adapt our user interface. Most importantly though, our work on cognitive monitoring is designed to provide clinical feedback to the elder. Based on the elder's preferences, he or she may choose to share this information with caregivers and clinicians.

Conclusion

We have demonstrated a software architecture for real-time unobtrusive monitoring of computer interactions for the purpose of inferring cognitive performance. This approach offers substantial benefits in being able to measure within subject changes over time in a natural setting. Our ability to detect trends in cognitive performance offers the possibility of early detection, both of future cognitive decline that could be treated early, and of near-term effects of medication interactions or more acute illnesses. Our early results demonstrate that elderly computer users enjoy the games and play them frequently. We also have early evidence of our metrics being able to distinguish between cognitively healthy elders and those with mild cognitive impairment. Our hope is that this monitoring information may be an inexpensive way of facilitating cognitive health management for elders, helping them maintain their quality of life and independence.

Acknowledgments

This work was supported by NIST ATP, the Intel Corporation, and by ORCATECH, Oregon's Roybal Center for Aging & Technology with a grant from NIA (P30AG024978).

References

[1] Garber AM. Long-term care, wealth, and health of the disabled elderly living in the community. In: Wise DA, ed. The Economics of Aging. Chicago: University of Chicago Press, 1989.

[2] Social Security Online, http://www.ssa.gov/OACT /TR/ TR04/ index.html, last update March 23, 2004.

[3] US Census Bureau, www.census.gov/population/socdemo / age.html, last revised: November 04, 2003.

[4] Kasper JD. Using the long-term care surveys: longitudinal and cross-sectional analyses of disabled older people. Proceedings of the 1987 Public Health Conference on Records and Statistics. DHHS Pub.No.88-1214. Hyattsville, MD: National Center for Health Statistics, 1988, pp. 353-358.

[5] Kiernan R, Meuller, J, Langston J, and Van Dyke C. The neurobehavioral cognitive status examination: a brief but differentiated approach to cognitive assessment. Annals of Internal Medicine 1987:107: 734-739.

[6] Petersen RC, Stevens JC, Ganguli M, Tangalos EG, Cummings JL, DeKosky ST. Practice parameter: early detection of dementia: Mild cognitive impairment (an evidence-based review): Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2001; 56(9):1133-42.

[7] Chen P, Ratcliff G, Phil D, Belle S, Cauley J, De Kosky S, and Ganguli M. Cognitive tests that best discriminate between presymptomatic AD and those who remain nondemented. Neurology 2000: 55: 1847-1853.

[8] Callahan CM, Hendrie HC, and Tierney WM. Documentation and evaluation of cognitive impairment in elderly primary care patients. American College of Physicians 1995; 122: 422-429.

[9] Mitrushina MN, Boone KB, and D'Elia DF. Handbook of Normative Data for Neuroopsychological Assessment. Oxford University Press, 1999.

[10] Howieson d, Bowman S, Coehlo D, Lim S, and Kaye J. Natural history of cognitive decline in the old old. Neurology 2003; 60: 1489-1494.

[11] Geldmacher DS and Hills EC. Effect of stimulus number, target-to-distractor ratio and motor speed on visuospatial search quality following traumatic brain injury. Brain Injury 1997; 11(1):59-66.

Address for correspondence

Holly B Jimison, PhD
Department of Medical Informatics and Clinical Epidemiology
BICC 507, Oregon Health & Science University
3181 SW Sam Jackson Park Rd
Portland, Oregon 97239 USA
jimisonh@ohsu.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Mobile Phone Computing for In-situ Cognitive Behavioral Therapy

Magnus Bang[a,b], Toomas Timpka[c], Henrik Eriksson[a], Einar Holm[d], Conny Nordin[e]

[a] *Department of Computer and Information Science, Linköping University, Sweden*
[b] *Interactive Institute, Stockholm, Sweden*
[c] *Department of Social Medicine and Public Health, Linköping University, Sweden*
[d] *Department Social and Economic Geography, Umeå University, Sweden*
[e] *Department of Neuroscience and Locomotion, Division of Psychiatry, Linköping University, Sweden*

Abstract

Cognitive behavioral therapy (CBT) for psychological disorders is becoming increasingly popular on the Internet. However, when using this workstation approach, components such as training and learning relaxation skills, problem solving, exposure exercises, and sleep management guidance must be done in the domestic environment. This paper describes design concepts for providing spatially explicit CBT with mobile phones. We reviewed and analyzed a set of treatment manuals to distinguish elements of CBT that can be improved and supported using mobile phone applications. The key advantage of mobile computing support in CBT is that multimedia can be applied to record, scale, and label anxiety-provoking situations where the need arises, which helps the CBT clients formulate and convey their thoughts and feelings to relatives and friends, as well as to therapists at subsequent treatment sessions.

Keywords:

mental health informatics, ubiquitous computing, cognitive behavioral therapy, social geography

Introduction

Psychiatric disorders are common in the industrialized world, which is illustrated by the observation that the cost of treating anxiety alone exceeds an estimated $40 billion annually in the United States [1]. Rehabilitation of individuals suffering from long-standing psychiatric and psychological disorders is often obstructed by insufficient coordination of the services provided by health care workers and public welfare agencies [2]. Thus there are strong incentives for developing self-support programs for this category of patients.

Cognitive behavioral therapy (CBT) is a widely acknowledged approach for treatment of numerous psychological disorders [3,4,5], and its efficacy has been demonstrated in several empirical studies and clinical trials [cf., 6,7].

CBT usually involves components such as motivation and education, relaxations skills, problem solving training, exposure exercises, and sleep management training. However, barriers related to accessibility and affordability have prevented many people from receiving this form of treatment [8]. In response to these problems, computer-mediated therapies have been introduced and used for more than a decade [9]. More recently, this strategy has also reached the Internet ("interapy") and definitely shows potential, particularly in terms of cost-effective distribution [10]. The online therapies typically contain self-help instructions as well as exercises that comprise either unstructured patient-therapist interactions or structured manual-based training with little, if any, therapist involvement [11,12].

A clear drawback of most of the current Internet-based therapies is that interaction is restricted to stationary computers. This may decrease the efficacy of training, cognitive restructuring, and learning, because the exercises become too abstract when performed in a decontextualized environment. Interestingly, a central tenet of CBT is self-managed recording of anxiety-provoking situations and exposure training conducted in real contexts. The basic hypothesis of our work is that, under appropriate circumstances, portable computer tools such as mobile phones have the potential to significantly improve both traditional CBT (*in situ*) and internet-based therapies.

As a step towards *using mobile phones as a component of in situ treatment,* we present a set of design requirements and concepts for 3G/4G mobile phones that emanate from an analysis of CBT manuals for generalized anxiety disorders (GADs) and information regarding the theory of social geography [13]. First, we outline the fundamental concepts of CBT, and thereafter we discuss our two approaches to design: ubiquitous computing and context-aware computing. We conclude our paper by discussing how modern mobile phone technologies like cameras and satellite positioning can be integrated into CBT self-help programs for clients with various psychiatric disorders such as agoraphobia and GAD.

Background

Cognitive behavioral therapy

CBT is used to treat mental disorders by attempting to modify negative belief systems and associated behaviors

[3,4,14]. The underlying concept is that distorted thoughts have a deleterious impact on emotions, and hence the objective of CBT sessions is to help people revise negative thinking and adopt new and healthier thoughts and actions. The adequacy of CBT for treating depression and generalized anxiety disorders is particularly well documented and acknowledged in the research community.

The components of CBT vary depending on the diagnosis and needs of the individual, but in all cases the general idea is to help the patient *identify and modify negative thinking patterns related to situations* that affect emotional responses. Counselors employ essentially three learning approaches in CBT: the process of *guided discovery using the Socratic method, thought diaries,* and *behavioral experiments*. Socratic questioning is a method to promote new alternative interpretations of situations. Once this process has started, the novel interpretations that emerge can be documented in a special diary [15] called a *dysfunctional thought record* (DTR). In the DTR, clients describe and keep track of common negative situations and the related emotions, automatic thoughts, and responses. Moreover, these qualities are scaled and outcomes of the situation are evaluated in light of the alternative thought patterns. The core idea is that the cognitive schema embedded in the diary will be internalized and automatically applied to situations in real life. The behavioral experiments in CBT aim to aid the clients in testing and validating held beliefs. For example, to treat people with a fear of wasps, counselors set up special experiments in which the phobia is put to the test by letting clients encounter the insects to ascertain whether the hypotheses are true (exposure therapy). Yet another important and integrated part of the treatment includes *reflective exercises and homework*. These assignments can even be done in the field, for instance to treat agoraphobia and social disorders.

Media in traditional CBT

Thus CBT is characterized by the use of a range of abstract methods and instruments such as the DTR, which are often accommodated in different types of technical artifacts. Tape recorders, pen and paper, thought diaries, activity schedules, and questionnaires are commonly used in CBT sessions and exercises [14]. Considering the example of management of phobias related to spiders and wasps, both models and living specimens of these arthropods are used in exposure treatment sessions. Counselors often give the client a ring binder at the onset of therapy to facilitate organization of materials as they are introduced sequentially during the treatment period. This binder contains the session outlines, the DTR, summaries, handouts, and homework instructions and exercises. Therefore, an important research question is how to replace the mentioned traditional tools with improved digital alternatives that preserve the cognitive schemata and their incorporated restructuring techniques.

Ubiquitous computing and mobile phones

Ubiquitous computing represents a shift from conventional desktop-based computing and solutions towards mobile computing and attempts to blend digitization into ordinary physical tools and environments [16]. The goal is to create a new class of user interfaces that more appropriately match our everyday activities and information needs and that do not require stationary work at a computer. A special area within this field is called context-aware computing [17]. According to Guanling and Kotz, mobile context-aware applications can identify and take advantage of contextual information in the physical environment, such as user location, time of day, nearby people and devices, and the user's activity [18]. Some examples of such applications are GPS-directed maps and special mobile games that utilize wireless Bluetooth as a proximity sensor to detect other gamers and active artifacts in the vicinity (social games).

Mobile phone applications are central components of ubiquitous and context-aware computing. Today, mobile phones are becoming increasingly sophisticated to include Internet access, cameras, web browsers, advanced media capabilities such as MP3, and features for GPS positioning (cf., Nokia N95). Some phones also have calendars and special fitness applications with motion sensors for use as a pedometer (cf., SonyEricsson W710i). Interestingly, it seems that mobile phones can provide suitable support for CBT treatment because they emphasize situatedness, mobility, thinking, and reflection.

Methods

We analyzed a set of manuals that are used by therapists to treat various psychiatric illnesses, particularly GADs [15,19, 20]. These manuals were chosen because our aim was to find general features of CBT to which mobile technology could be applied as a means of improving the traditional treatment components. We also used Hägerstrand's framework outlining 'social pockets of local order' as a foundation for provision of CBT *in situ* [13]. This framework departs from each person's residence and displays the different meeting places as 'social pockets'. With this framework, anxiety patients can be represented and followed in terms of, for example, movement patterns and changes in those patterns over time, as well as comparisons with normal populations. Below we describe how mobile phones can be used to specifically deliver CBT in this context.

Results

The digital dysfunctional thought record

The digital dysfunctional thought record (DDTR) is an electronic version of Beck's empirically validated DTR for recording negative situations and the related emotions and automatic thoughts on paper.

Figure 1 - Clients use the DDTR to record situations in which anxiety arises

Figure 1 shows a prototype of a DDTR for mobile phones that we implemented in Macromedia Flash Lite. Our version employs the camera of the mobile phone to document situations and places (tagged by geocode) that are experienced as unpleasant. It is also possible to add both text and audio input to comment on the pictures and short films. Scaling the anxiety level of the specific situation and media is done using the wheel on the phone (Beck scale 1–100).

A well-known problem with the CBT exercises and anxiety recording is that many patients do not have labels for their emotions, and they find it difficult to express their thoughts [14]. Researchers have suggested that clients should use *one word* to denote each type of feeling [21]. We support this approach in our application providing basic lists of definitions and synonyms for feelings that can be used during recording. From the perspective of CBT, there are some intrinsic problems with this method, which are discussed in the following section. In addition, it is imperative that the content of the DDTR can be shared with the therapist before each CBT session. Storing the media documentation as a single encrypted file makes transfer over the Internet convenient and secure. Naturally, an alternative is to transfer the files wirelessly via Bluetooth at the start of the face-to-face session. Nevertheless, we believe the main benefit of recording real situations and places for CBT is that it facilitates recall of negative thoughts, which makes it easier for the client to convey problems and situations to the therapist during the actual session.

Relaxation and sleep training

Learning relaxation techniques can benefit many patients, particularly those suffering from anxiety disorders such as a GAD. Relaxation training can reduce physiological arousal, and cognitive procedures can modify excessive worry. Basically, there are two relaxation techniques that are used to deal with anxiety in CBT: breathing training and progressive muscle relaxation [20, 22]. These components are often introduced before the actual cognitive restructuring sessions. Programs for training relaxation skills via the mobile phone are promising. For example, a therapist can give verbal instructions on how to breathe and record them on audiotape or a CD for later use by the client. It would be a relatively straightforward task to transfer this information to a mobile phone equipped with an MP3 player. There are several benefits associated with having the relaxation media pervasively available to allow meditation wherever the client prefers.

Supporting daily routines

Many psychological disorders are associated with reduced organizational skills and problems with memory recall, as a combined result of the illness and the medication. This problem leads to difficulties in managing normal events and even in remembering the visits to the therapist. Mobile phone technologies include several aids (e.g., calendars that have alarms and provide reminders) that can be available everywhere and can help CBT patients structure their days. These applications are not currently applied in CBT, but they can easily be introduced for that purpose.

Social and spatial isolation are also connected with psychological disorders. According to the theory of social pockets, activity patterns and spatial movement from home are indicators of the severity of illness. Context-aware applications that track movements in space and also the number of calls and messages that are made could constitute a useful scale for implicitly tracking progress. Special phones with step-counter and GPS motion-tracking software are already on the market. These phones measure walking distance and speed and calories burned during the day, and thus they can be used to develop an application that combines motion sensing and a calendar with alarms to aid patients with dysfunctional activity patterns. For example, a context-aware application could be designed to monitor sleeping and activity patterns and possibly also to enforce daily routines and—in cooperation with the counselor—promote a more active lifestyle.

Distraction support

Distraction techniques are often used in CBT to take a patient's mind off his/her problems and thereby lower anxiety levels. Physical activity, mental imaging, and doing arithmetic are ways of blocking the focus on feelings and directing the patient's thoughts to real-world activities and tasks. Special media such as images and certain interactive games can obviously be incorporated into the phone to offer a distracting effect. However, many therapists feel that such techniques should only be applied solely as a first-aid approach, because they can be used to avoid emotions and negative thoughts, which can have an impact on the therapeutic outcome.

Therefore, before integrating such techniques into mobile applications, studies should be conducted to evaluate short and long-term effects.

Session recording

According to Willis and Sanders [14], audio recording of counseling sessions is an invaluable strategy. In short,

giving clients the means to listen to previous sessions is helpful, because it reinforces what has been discussed and gives client useful feedback (memory). We feel that future CBT applications for mobile phones should include such basic recording software to allow patients to listen to their therapy sessions at any time and anywhere.

Discussion

Mobile digital technologies can be applied to supply spatially explicit CBT *in situ*. We believe that an appropriate starting point is to support key practices in CBT. For instance, mobile phones can help patients by making it possible to record and spatially and temporally organize anxiety-provoking situations, and also to facilitate the labeling of thoughts and feelings. As mentioned above, cognitive restructuring is a basic principle of CBT and the traditional tools such as the DTR has been developed specifically to support this. The idea is that the CBT clients will internalize these schemas and automatically apply them even after conclusion of the CBT period and throughout life. Thus, the question arises as to whether new digital versions of the CBT tools, such as our DDTR, encompass the necessary cognitive properties. The traditional diaries and scales used to record situations, as well as the schemata they embody for cognitive restructuring, have been empirically tested for decades. Therefore, efforts must be made to validate the mobile digital CBT tools by applying the same procedures used to test the traditional CBT tools.

Our DDTR has a dictionary of terms for labeling feelings, and this is included to assist users in finding appropriate words to describe and audio record their experiences. However, it is possible that this dictionary can inappropriately guide the user's choice of words. If that is the case, it could lead to false labeling, which would naturally impair the therapy. Clearly, more research is needed in this area.

We are currently developing a set of phone-based applications to support the practice of CBT, and we are also performing a small study to address the above-mentioned question regarding directed labeling. We are conducting that work from the perspective of participatory design, so that the CBT clients themselves can influence the construction of the applications.

Conclusion

In this paper, we outline a set of tools that can be part of a client mobile multimedia application to support CBT. Specifically, we describe a new way to document and label the anxiety-provoking situations that are part of regular CBT treatment. The diaries that are traditionally used to record situations, as well as the cognitive restructuring schemata they embody, represent evidence-based clinical technologies. Moving to modern media in CBT, as we suggest here, will require comprehensive evaluation. Nonetheless, we believe that new media can have a profound impact on CBT, in both a positive and a negative sense. Thus we strongly recommend that implementation of novel applications and designs for mobile CBT should be accompanied by controlled trials that assess the efficacy of these innovative strategies.

References

[1] Greenberg PE, Sisitsky T, Kessler RC, Finkelstein SN, Berndt ER, Davidson JR, Ballenger JC, Fyer AJ. The economic burden of anxiety disorders in the 1990s. J Clin Psychiatry 1999 60(7):427-35.

[2] Timpka T, Bang M, Delblanco T, Walker J. Information infrastructure for inter-organizational mental health services: An actor network theory analysis of psychiatric rehabilitation. Journal of Biomedical Informatics. In Press.

[3] Beck AT (ed). Cognitive therapy of depression. New York: Guilford Press 1979.

[4] Beck J. The past and future of cognitive therapy. Psychother Pract Res.1997; 6: 276-284.

[5] Myhr G, Payne K. Cost-effectiveness of cognitive-behavioural therapy for mental disorders: implications for public health care funding policy in Canada. Can J Psychiatry 2006;10:662-70.

[6] Mohman J, Gorenstein EE, Kleber M, de Jesus M, Gorman JM, Papp JA. Standard and enhanced cognitive behavior therapy for late-life generalized anxiety disorder: Two pilot investigations. American Journal of Geriatric Psychiatry 2004;11:24-32.

[7] Brown TA, O'Leary TA, Barlow DH. Generalized anxiety disorder. In: Barlow DH, Ed. Clinical handbook of psychological disorders. NewYork: Guilford, 2001.

[8] National Institute of Health. *Treatment of panic disorder.* Paper presented at the NIH Consensus development conference consensus statement (Volume 9), Bethesda, MD.

[9] Marks I, Shaw S, Parkin R. Computer-aided treatments of mental health problems. Clinical Psychology: Science and Practice 1998;5:151-170.

[10] Lange A, Schrieken B, van de Ven JP, Bredeweg B, Emmelkamp PM, van der Kolk, Lydsdottir L, Massaro, M, Reuvers A. "Interapy": The effects of a short protocolled treatment of posttraumatic stress and pathological grief through the internet. Behavioural and Cognitive Psychotherapy 2000;28:175-192.

[11] Carlbring P, Westling BE, Ljungstrand P, Ekselius L, Andersson G. Treatment of Panic Disorder Via the Internet: A Randomized Trial of a Self-Help Program. Behaviour Therapy 2001; 32(4):751-764.

[12] Andersson G, Carlbring P. Editorial – special issue: Internet and cognitive behaviour therapy: new opportunities for treatment and assessment. Cognitive Behaviour Therapy 2003;32:97–99.

[13] Hagerstrand T. Time-geography: focus on the corporeality of men, society and environment. In: The Science and Praxis of Complexity. New York: United Nations University, 1985.

[14] Willis F, Sanders D. Cognitive Therapy: Transforming the Image. London: Sage Publications, 1997.

[15] Beck JS. Cognitive Therapy: Basics and Beyond. New York: Guilford Press 1995.

[16] Weiser M. Hot Topics: Ubiquitous Computing. IEEE Computer, October 1993.

[17] Shilit BN., Adams, NI, Want R. Context-Aware Computing Applications. In: Proceedings of the Workshop on Mobile Computing Systems and Applications. Santa Cruz: IEEE Computer Society 1994: 85-90.

[18] Guanling C, Kotz D. A Survey of Context-Aware Mobile Computing Research. Technical Report TR2000-381, Department of Computer Science, Dartmouth College, 2000.

[19] Dick LP, Gallagher-Thompson D, Coon DW. Cognitive-Behavioral Therapy for Late-Life Depression: A therapist manual. Palo Alto, CA: Older Adult and Family Center, Veterans Affairs Palo Alto Health Care Center, 1995.
[20] Stanley MA, Diefenbach GJ, Hopko DR. Cognitive behavioral treatment for older adults with generalized anxiety disorder. Behavior Modification 2004, 28 (1), 73-117.
[21] Greenberger D, Padesky C. Mind Over Mood. New York Guilford Press 1995.
[22] Bernstein, DA, Borkovec TD. Progressive relaxation training. Champaign, IL: Research Press 1973.

Acknowledgments

This research was supported by a grant from the Health Informatics program at VINNOVA.

Address for correspondence

Dr. Magnus Bang, E-mail: magba@ida.liu.se
Department of Computer and Information Science
Linkoping University, SE-581 83 Linkoping, Sweden

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Methods for Measuring the Impact of Health Information Technologies on Clinicians' Patterns of Work and Communication

Johanna I Westbrook[a], Amanda Ampt[a], Margaret Williamson[b], Ken Nguyen[a] Leanne Kearney[a]

[a]*Health Informatics Research & Evaluation Unit, Faculty of Health Sciences, University of Sydney, Australia ,*
[b] *National Prescribing Service, Australia*

Abstract

Evidence regarding how health information technologies influence clinical work patterns and support efficient practices is limited. Traditional paper-based data collection methods are unable to capture clinical work complexity and communication patterns. Our objective was to design and test an electronic data collection tool for work measurement studies which would allow efficient, accurate and reliable data collection, and capture work complexity.

We developed software on a personal digital assistant (PDA) which captures details of nurses' work; what task, with whom, and with what; multi-tasking; interruptions and event duration. During field-testing over seven months across four hospital wards, fifty-two nurses were observed for 250 hours. Inter-rater reliability scores were maintained at over 85%. Only 1% of tasks did not match the classification developed. Over 40% of nurses' time was spent in direct care or professional communication, with 11.8% in multi-tasking. Nurses were interrupted approximately every 49 minutes. One quarter of interruptions occurred while nurses were preparing or administering medications. This approach produces data which provides greater insights into patterns of clinician's work than has previously been possible.

Keywords:

computerised order entry systems; nursing staff; medication systems; time and motion studies; computers, handheld; observation; interdisciplinary communication

Introduction

Evidence regarding how health information technologies influence patterns of clinical work and support efficient work practices is limited. A recent systematic review [1]uncovered 23 studies since 1984 which examined the impact of system use on clinicians' (doctor and nurses') time. These studies in general adopted either work sampling or time and motion methods. Only six (26%) examined work on general wards, all in US hospitals, while the remainder focused on specialized settings (e.g., ICU and general practice). Overall, studies which compared electronic with paper systems and calculated task time per patient or consultation, reported that computer use increased time required to complete tasks. Studies which examined task time across multiple patients or working shifts found computer use was more time efficient than paper-based systems [1].

Studies of changes in work distribution and communication patterns following system use are less prevalent, but do include evidence of changes. For example in the US, where nurses usually transcribe handwritten medication orders, CPOE eliminates this task [2, 3]. Following the introduction of a CPOE system clinicians may sequence work differently. As Callen et al [4] found, clinicians reported " ..if you are waiting for something on the computer you go and do something else." (p648) This may result in, for example, batching the ordering of patients' tests to one time of the day. Shu et. al. [5] found interns in a US hospital spent more time alone and less time with other doctors after system introduction. A French hospital study [6] of doctor-nurse communication around medications showed that a CPOE system, in comparison with paper-based medication records, resulted in a move from synchronous communication to asynchronous. This introduced opportunities for misunderstandings and increased the extent to which nurses had to make assumptions about orders. Carpenter et. al. [7] also reported a tendency for doctors in a US hospital to talk with nurses less about medication orders following system implementation.

Understanding these shifts in patterns of communication between clinicians are important as poor communication wastes time, threatens patient care and may be one of the chief culprits behind preventable adverse events in clinical practice [8]. Any potential negative consequences of changes in communication patterns may be more than offset by the improvements in information exchange provided by having legible, easily accessible information which computerized systems afford clinicians. However until we have better quality data about how systems disrupt existing patterns of clinical work and communication we cannot move to re-design work practices or systems in ways which avoid any possible negative outcomes.

This research agenda needs to continue to progress beyond answering the question on whether the use of a computer save a clinician time to questions about how patterns of work are re-arranged in response to the introduction of health technologies. Where time is released, or additional time consumed, how do clinicians re-distribute their time amongst work tasks? What amount of variation exists among different clinical sub-groups and do work tasks get

Selected for best paper award.

re-distributed across groups? For example, if senior clinicians are found to spend less time in patient ordering following computerisation, is this because the system is efficient or because they have re-allocated this task, either explicitly or implicitly, to their junior colleagues? We need to examine how system use interferes with communication processes and as Gorman et. al. [9] suggest, ensure that such systems *".. facilitate care without interfering with or eliminating aspects of the process that are essential to high reliability performance in the face of urgency, uncertainty and interruptions"* (p383)

We require studies which investigate whether changes in patterns of clinical work result in improved care delivery, patient outcomes and the work experiences of health professionals. These questions require a multi-method approach, and work measurement studies form an important component of such investigation. Researchers should be able to build upon previous work undertaken in different settings and countries. This requires standardization of measurement approaches and the adoption of valid and reliable measurement tools. Major factors identified for the paucity of evidence in this area are the limitations and varieties of methods used [1, 10, 11], the difficulties of capturing the non-linear and interruptive nature of clinical work, and the lack of consistency in the application of rigorous research methods.

Our objective was to design, build and test an electronic data collection tool for use in work measurement studies which would allow efficient, accurate and reliable data collection while also capturing a greater level of work complexity than previous paper-based methods have allowed. Results from a small number of previous studies [12, 13] suggest that the use of handheld computers (personal digital assistants – PDAs) may be useful for this task but researchers have presented minimal information about application and reliability issues relating to these tools. Building upon our previous work designing a paper-based, multi-dimensional work classification tool for nurses [11], we sought to investigate how much additional detail and task complexity data we could add using a PDA without reducing data accuracy or reliability. In this paper we detail the development of these methods and tool, outline the data collection process using the PDA in an observational study of fifty-two nurses undertaking everyday work activities over 250 hours in wards of an academic hospital, and provide a summary of the results. This comprised the first stage in a study to measure the impact of a commercial electronic medication management system (e-MMS) on doctors' and nurses' work and communication patterns.

Using these baseline data we aimed to examine how nurses distributed their time across major work tasks and communication events, and how patterns vary, for example by nurse classification. We also sought to measure the extent of multi-tasking and interruptions to nurses' work.

Materials and methods

Design of a multi-dimensional work task classification

Our first objective was to develop a work task classification system which would be incorporated in the PDA data collection tool. As a basis for this we used a multi-dimensional work measurement classification which we had applied in a paper-based work-sampling study of nurses and showed high levels of inter-rater reliability and face validity [11]. We extended the classification to contain greater detail about work tasks which previous literature indicated may be most susceptible to change following system introduction and thus would address our long-term aim to use the PDA in a study to examine the impact of an e-MMS on clinicians' patterns of work and communication.

The classification captures the complexity of nursing tasks in terms of: i) the activity being undertaken, ii) other people involved in the task, and iii) the method of task performance. The ten broad categories of work tasks and definitions are listed in Table 1.

Table 1 - Broad work task categories

Work Task	Definition
Direct Care	Tasks directly involved with patient care, eg direct communication with pt. &/or family, bathing, applying dressings etc.
Indirect Care	All tasks indirectly related to patient care (eg reviewing results, planning care)
Medication Tasks	All tasks associated with medication, includes preparation, administration, documentation, discussion & clarification
Documentation	Documentation (paper and electronic), excludes medication documentation
Professional Communication	All non-medication related communication with another health professional, includes ward & patient hand overs
Ward Related Activities	Ward activities, includes coordinating beds & staffing
In Transit	Time between tasks and between patients
Supervision	Supervising others, including students
Social	All non work communication, meal breaks
Other	Any other task not included above

Using a PDA allowed us to add further detail to some work tasks by using drop down submenus. Figure 1 shows the 10 broad work task categories and Figure 2 one of the submenus for medication-related tasks. Further we wished to capture instances of multi-tasking (undertaking two tasks at the same time) and interruptions. This had not been feasible with a paper version. Interruptions were defined as the ceasing of one task in order to attend to another task. For example, a nurse who stops administering a drug in order to answer a question from a colleague. If the nurse

continues to administer the medication while answering her/his colleague the activity would be recorded as multi-tasking.

Data collection tool

We designed a work measurement data collection tool using an HP iPAQ rx3000 Pocket PC running Windows Mobile 2003 with a SD card expansion slot. The PDA program was developed in Visual C# using the .NET Compact Framework v1.0. Using ActiveSync enabled the use of SQL and the synchronization of multiple PDA databases with one central database on a PC. Figure 1 shows the interface design with the 10 broad mutually exclusive work task categories listed on the left. Each task is automatically time-stamped. On the top right hand side the data collector records who else is involved in the task, and in the bottom right of the screen any tools/equipment that are used. Selection from any one of the three sections will immediately start a new time stamp – thus if the data collector is aware a new activity has started as the nurse is now with another professional, but is unsure what the activity is on initial observation, she/he can select another entry point, such as who the nurse is with which might be easier to distinguish.

If the observed subject is interrupted or is multi-tasking, the observer hits the "Add" button. If the subject is multi-tasking then the observer will continue to record all work tasks until the multi-tasking ends (End Multi). If the subject is interrupted then the interruption button is selected and the interrupted task appears as a tag on the bottom of the screen. When the subject returns to the task the observer is able to pick up this pending task. This allows, for example, calculation of the time from interruption until final completion of the task. The 'Ignore' button is used as a quick way for the observer to delete data recorded in error.

All recorded data are placed in a database which is stored on the SD card. At the end of an observation session the

Figure 1 - PDA work measurement data collection tool

PDA is synchronised with a computer running Microsoft SQL Server using merge replication. SQL Server replication is commonly described by using the publisher/subscriber metaphor. A database server that makes data available for replication (source server) is referred to as the publisher; a collection of one or more database objects that are enabled for replication is called a publication.

Figure 2 - Drop down submenu displayed

One or more servers that receive data and/or transactions from the publisher are called subscribers (the PDAs, in this case). Replication is managed by the system database, which by default is called distribution. A distribution database—which can reside on the publisher, subscriber, or on a separate server or computer—is created when replication is configured.

The server that hosts the distribution database is referred to as the distribution server or distributor. Merge replication combines data from multiple sources into a single central database. This allows multiple observers to use different PDAs for data collection. On synchronization, data from each PDA is then transferred to the central database.

An extraction program was created to extract the data from the distributor. This was written in Visual C# using the .NET Framework v1.1. This allowed data to be imported into a statistical package such as SPSS for analysis. The data collection tool, synchronization and extraction programs underwent considerable field trials. Once these were completed we undertook a full-scale study to examine nurses' patterns of work and communication tasks. This constituted the first stage in a study to investigate changes in nurses' work patterns following the implementation of a commercial electronic medication management system.

The work measurement study was conducted over seven months (July 2005-February 2006) in four wards (respiratory, renal/vascular and two geriatric), at a major academic teaching hospital in Sydney, Australia. Data were collected between the hours of 7am – 7pm, Monday to Fridays.

Procedures and participants

Nurses on the four wards were invited to participate at ward information sessions. Following signed consent nurses were assigned a study identification number, and demographic information regarding their age, nurse classification (e.g., enrolled nurse, registered nurse-new graduate, registered nurse, clinical nurses specialist etc.), and length of experience was collected. The observer shadowed nurse participants for an average of one hour blocks, noting all work tasks performed using the PDA to record data. When the participant nurse engaged with patients, visitors, or other health professionals, the nurse was asked to introduce the observer and seek permission to continue. Alternatively the observer would identify themselves. The

study was approved by the ethics committees of the hospital and the University of NSW.

Sample size calculations, based upon our previous work-sampling studies [11],were made which determined the number and times of hours of observation per nurse category required. Our intention was to be able to detect differences in the broad work categories pre and post system implementation. As each category had its own mean and standard deviation, it was decided to use medication-related tasks to determine the sample size – this category had the largest standard deviation and is the most directly related to the proposed system change. We determined that at alpha=0.05 and with power of 80%, 226 hours of observation were required to detect a three minute difference in the proportion of time spent in medication-related tasks pre and post system introduction using a two-tailed t-test. We achieved 250 hours of data collection during the pre-implementation phase.

Data collector training

All observers were clinically experienced registered nurses. The initial training was delivered by one of the authors (AA), who had experience with previous work measurement studies [9]. Areas covered included outline of study purpose and methodology, instruction in PDA use with explanation of definitions and practice scenarios, introduction to study wards, practice sessions on the wards followed by inter-rater reliability testing. Once the second data collector was trained (LK), she then delivered all subsequent training to three more data collectors. Approximately fifteen hours was required for each data collector to be trained.

Inter-rater reliability and data analysis

Following training of each new data collector, inter-rater reliability tests were conducted, as well as at other random times. Two data collectors simultaneously but independently observed a nurse and recorded work tasks on their own PDA for approximately 45 minutes. Data were then analysed and compared. New collectors did not start collecting data until overall percentage time in tasks was in agreement with a more experienced collector over 85% of the time. All data collectors maintained this level of agreement, (range 85%-98%).

Descriptive statistics with 95% confidence intervals were calculated. For this paper, differences in task time distribution by wards and nurse classification were analysed using SPSS.

Results

Work task distribution

Fifty-two nurses over four wards were observed for a total of 250 hours and 20 minutes. During this time 15,533 tasks were recorded. Figure 3 shows the distribution of observed time spent in different work tasks. Direct care and professional communication accounted for 41% of total task time.

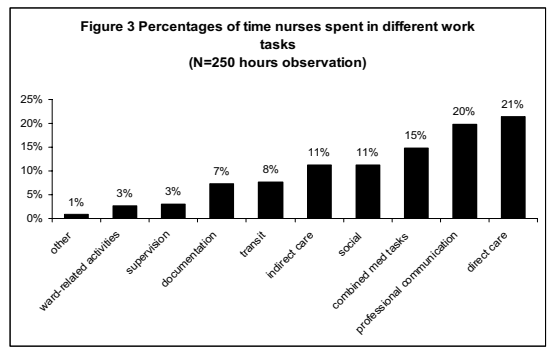


Figure 3 - Percentages of time nurses spent in different work tasks

The pattern of task time distribution differed between enrolled nurses and registered nurses. As Figure 4 shows enrolled nurses spent a greater proportion of their time in direct care activities while registered nurses spent comparatively more time in medication tasks and professional communication.

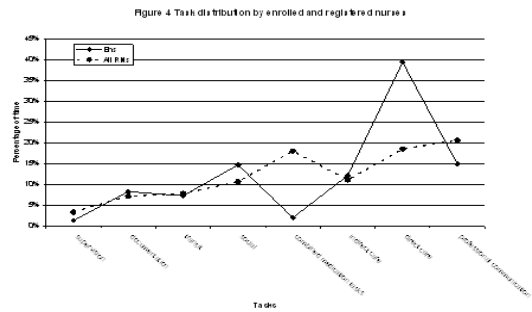


Figure 4 - Task distribution by enrolled and registered nurse

We examined task time distribution by wards and found considerable similarity in the pattern of time distribution despite differences in the ward specialties. As Figure 5 shows, for nearly all task categories the 95% confidence intervals overlapped.

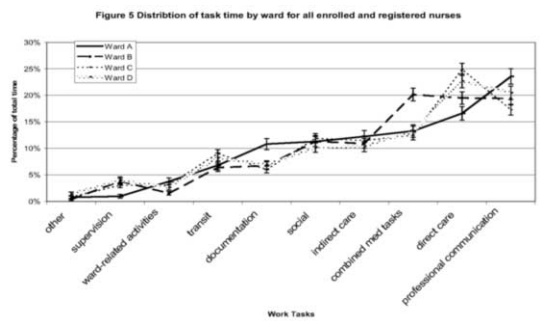


Figure 5 - Distribution of task time by ward for all enrolled and registered nurses

Multi-tasking and interruptions

The total task time recorded was 279 hours 51 minutes, compared to 250 hours 21 minutes of observed time, which revealed that nurses spent 11.8% (29hrs 30 minutes) of observed time multi-tasking in which two or more tasks were performed in parallel. On average nurses were found to multi-task for approximately one hour and two minutes every shift. In addition, on average, nurses were interrupted every 49 minutes. One quarter of all interruptions occurred while nurses were undertaking medication-related tasks and a further 23% of interruptions occurred while nurses were documenting.

Professional communication

Overall nurses spent 20% of observed time in professional communication which equates to approx. 12 minutes/ hour. The majority (77%) of this time was spent communicating with other nurses. Only 8%, a mean of 8 minutes per shift (of which 45mins is meal break) was spent talking with doctors.

Discussion

Our aim was to design research methods and a tool to allow efficient, reliable and accurate data collection about clinicians' patterns of work and communication, which could potentially be used across specialties, professional groups, settings and countries. The observational methods and the PDA data collection tool proved to be a reliable means for collecting complex, multi-dimensional data about nurses' work and communication patterns. High levels of inter-rater reliability were achieved with an average of 15 hours of training. Training, recruitment and 250 hours of data collection were undertaken by the 1.5 full-time equivalent staff over approximately 10 months. Thus, we found it to be a relatively time-efficient method, particularly as data are entered electronically in real-time reducing transcription errors. The work task classification and accompanying definitions dealt well with the array of nurses' observed work tasks with less than 1% of all tasks being categorized in the other category.

Our initial analysis comparing ENs and RNs provides some validation of the data collected. For example, ENs are legally restricted in their performance of many medication-related tasks and this was demonstrated in the data. Further examination of work tasks by nurse classification will provide a useful means in the second stage of the research by which to examine how tasks may be redistributed across nursing classifications following the introduction of the electronic medication management system.

The results demonstrated the greater detail of data about work and communication patterns which can be obtained using the PDA tool compared to traditional paper-based methods. A good example was new data about multi-tasking and interruptions. Most previous studies of interruptions to clinical work have not distinguished between interruptions and multi-tasking. This tool provides a useful way of doing this.

Further, more detailed analyses, such as which tasks are interrupted and which tasks are more likely to be interruptors are possible. Interestingly we found that the greatest proportion of interruptions occurred while nurses were undertaking medication tasks. Such a finding raises the question of the role of interruptions in contributing to medication preparation and administration errors. We are currently investigating this issue further. The data cannot reveal whether interruptions had negative or positive effects. For example, the outcome of an interruption to stop administering a drug in error versus an interruption which occurs in the middle of a drug dose calculation will be quite different.

The PDA allows for detailed capture of patterns of work tasks and communication. This information, along with identification of tasks and situations that may lead to cognitive overload through multi-tasking and interruption, is essential in understanding how clinical information systems may impact upon the quality and safety of care.

Acknowledgments

J Forbes, A McLaren, K Gatwood for assistance with data collection. This research was funded by an ARC grant in partnership with NSW Health. JW is supported by an NHMRC Fellowship. We thank the hospital & nurses who participated in the study.

References

[1] Poissant L, Pereira J, Tamblyn R, Kawasumi Y. The impact of electronic health records on time efficiency of physicians and nurses: A systematic review. Journal of American Medical Informatics Association 2005;12(5):505-516.

[2] Ball MJ, Weaver C, Abbott PA. Enabling technologies promise to revitalize the role of nursing in an era of patient safety. International Journal of Medical Informatics 2003;69(1):29-38.

[3] Kaushal R, Jha A, Franz C, Glaser J, Shetty K, Jaggi T, et al. Return on investmant for a computerized physician order entry system. Journal of American Medical Informatics Association 2006;13(3):261-266.

[4] Callen J. The people and organisational issues which impact on the diffusion of IT systems in healthcare facilities [PhD]. Sydney: University of New South Wales; 2006.

[5] Shu K, Boyle D, Spurr C, Horsky J, Heiman H, O'Connor P, et al. Comparison of time spent writing orders on paper with computerized physician order entry. In: Patel Vea, editor. Medinfo; 2001; London: IOS Press; 2001. p. 1207-1211.

[6] Beuscart-Zephir M, Pelayo S, Anceaux F, Meaux J, Degroisse M, Degoulet P. Impact of CPOE on doctor-nurse cooperation for the medication ordering and administration process. International Journal of Medical Informatics 2005;74:629-641.

[7] Carpenter J, Gorman P. What's so special about medications: A pharamcist's observations from the POE study. In: Proceedings of American Medical Informatics Association Symposium 2001; 2001; 2001.

[8] Gosbee J. Communication among health professionals. British Medical Journal 1998;316:642.

[9] Gorman P, Lavelle M, Ash J. Order creation and communication in healthcare. Methods of Information in Medicine 2003;42:376-384.

[10] Burke T, McKee J, et al. A comparison of time-and-motion and self-reporting methods of work measurement. Journal of Nursing Administration 2000;30(3):118-125.

[11] Ampt A, Westbrook J, Creswick N, Mallock N. Comparison of self-reported and observational work sampling techniques to measure nurses' work. Journal of Health Services Research and Policy 2007;In press.

[12] Asaro P. Sychronized time-motion study in the emergency department using a handheld computer application. In: Proceedings of the World Congress on Medical Informatics (MedInfo); 2004; San Francisco: ISO Press; 2004. p. 701-705.

[13] Starren J, Chan S, Tahil F, White T. When seconds are counted: Tools formobile, high resolution time-motion studies. In: Proceedings of American Medical Informatics Symposium; 2000; 2000. p. 833-837.

Address for correspondence

Professor Johanna I Westbrook
email - J.Westbrook@usyd.edu.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Enhancing User Acceptance of Mandated Mobile Health Information Systems: The *e*POC (*electronic* Point-Of-Care Project) Experience

Lois Burgess[a], Jason Sargent[a]

[a] *Centre for Health Service Development, University of Wollongong, Australia*

Abstract

From a clinical perspective, the use of mobile technologies, such as Personal Digital Assistants (PDAs) within hospital environments is not new. A paradigm shift however is underway towards the acceptance and utility of these systems within mobile-based healthcare environments. Introducing new technologies and associated work practices has intrinsic risks which must be addressed. This paper contends that intervening to address user concerns as they arise throughout the system development lifecycle will lead to greater levels of user acceptance, while ultimately enhancing the deliverability of a system that provides a best fit with end user needs. It is envisaged this research will lead to the development of a formalised user acceptance framework based on an agile approach to user acceptance measurement. The results of an ongoing study of user perceptions towards a mandated electronic point-of-care information system in the Northern Illawarra Ambulatory Care Team (TACT) are presented.

Keywords:

ambulatory care, handheld computers, point-of-care, information systems, user acceptance

Introduction

A considered, measured transition from Health to eHealth is currently underway. While the move towards an electronic health record (EHR) is often the focus of discussion in this regard, it is the myriad in-situ electronic clinical information systems under development in which [The] effective use of Information Management and Technology and the implementation of core standards, applications and support networking will promote an environment where necessary information about a patient will be available to clinicians when and where it is needed [1]. For small, site-specific IS development projects to leverage on the advantages that this presents, the authors of this paper contend that agile systems development methodologies should be engaged from project initiation to enhance end-user acceptance of the delivered system. The decision to do so enables a best fit system for the end-user (clinician), the provider, and most importantly the patient. This is particularly important for mandated technology implementations. In scenarios of mandated technology implementation, impressions may exist whereby any form of consultative input from end-users is inherently removed; leaving end-users disillusioned with the technology. Moreover, where the intended end-users of a system are traditionally averse to technology through entrenched paper-based work practices (such as community health workers) the process of managing system change bears a considerable determination in system implementation success. This paper outlines an agile approach to end-user acceptance, taken to proactively address issues surrounding transition from traditional paper-based systems to mandated electronic systems in a community health setting with historically low levels of technology acceptance and use. The ePOC (electronic point-of-care) project is used to demonstrate how higher levels of user acceptance of mobile health information systems are attainable, even in mandated implementation environments.

The *electronic* Point of Care Project

The *electronic* Point-of-Care Personal Digital Assistant project (ePOC) is a multi-phase, collaborative research and development, Australian Research Council 3yr funded project, comprising a research team from the University of Wollongong, Flinders University, the University of South Australia and Pen Computer Systems Pty Ltd. The project client is The Ambulatory Care Team (TACT) Northern Illawarra which is supported by South Eastern Sydney and Illawarra Area Health Service (SESIAHS).

TACT employ 21 staff members consisting of 3 Doctors (including a Medical Director), 1 Nurse Unit Manager, 13 nurses (4 full-time and 9 Part-time), 2 Pharmacists (part time), 1 Physiotherapist and 1 COPD coordinator[1]). TACT Northern Illawarra[2] provide outpatient healthcare delivery; where patients are given a choice of having treatment in their usual place of residence (including aged care facilities) or other locations as an alternative to hospital. Community-based health services within New South Wales, Australia deliver over eight million occasions of service per annum provided by more than 7,000 clinicians from more than eight hundred and fifty health service loca-

1 Chronic Obstructive Pulmonary Disease Specialist.
2 Northern Illawarra comprises Wollongong, Shellharbour and Kiama Local Government Areas (LGAs)

tions. The cost of this service is estimated to be almost $450 Million dollars per annum [2].

In essence, the ePOC project is developing an integrated Ambulatory Care information system, deployed on a Personal Digital Assistant (PDA) platform. A PDA based point-of-care system for TACT is significant in that it will provide for *collection*, *delivery and exchange* of timely information (both text and images) at the point-of-care leading to a more efficient health care system. Current systems at TACT are paper-based and are limited to what the healthcare worker can effectively carry. The key advantages of a PDA system are its high mobility and flexibility in matching complex healthcare workflow requirements as well as immediate updating of healthcare records at the point-of-care. PDAs as wireless deployment platforms for mobile-based *hospital* clinical information have proven to be among the most cost effective ways to improve patient care quality and reduce medical data collection errors. Medical professionals empowered by information make better decisions while at the patient's bedside [3]. The challenge now is to extend information empowerment of clinicians from a patients (hospital) bedside to (home) point-of-care. It is within this project context that a new approach to technology acceptance, aimed at enhancing user acceptance of mobile health systems, utilising concepts drawn from agile systems development methodologies has been developed and is currently being tested and validated.

Theoretical basis of the study

The theoretical basis of this study centres on Dynamic Systems Development Methodology (DSDM) [4] and the Unified Theory of Acceptance and Use of Technology (UTAUT) [5]. From the outset, the research approach adopted by the ePOC project has centered on a consultative, open approach with all project team members; in particular, project team health area managers (Clinical, Research and IT) and the intended end users (TACT Doctors, Nurses and Para-Health Professionals) of the PDA application. To facilitate this approach, the ePOC project incorporates a number of research and development phases. According to Davis and Venkatesh, taking an iterative, consultative approach helps focus upon the identification, correction and prevention of requirements errors that have been introduced in the original specification of requirements [ibid].

The Dynamic Systems Development Methodology (DSDM)

Dynamic Systems Development Methodology is a framework supported by its continuous user involvement in an iterative development and incremental approach which is responsive to changing requirements, in order to develop a system that meets the business needs on time and on budget. It is one of a number of agile methods for developing software and forms part of the Agile Alliance. DSDM was developed in the United Kingdom in the 1990s by a consortium of vendors and experts in the field of Information System (IS) development, the DSDM Consortium. [6]

There are nine principles of DSDM [ibid]. While all nine are noteworthy, this paper focuses on the following six:

- Active user involvement.
- A focus on frequent delivery of products.
- Iterative and incremental development to ensure convergence on an accurate business solution.
- Reversible changes during development.
- Integrated testing throughout the life cycle.
- Collaboration and cooperation between all stakeholders.

The Unified Theory of Acceptance and Use of Technology (UTAT)

The Unified Theory of Acceptance and Use of Technology (UTAUT) was developed through a review and consolidation of the constructs of eight models that earlier research had employed to explain IS usage behaviour (theory of reasoned action, technology acceptance model, motivational model, theory of planned behavior, a combined theory of planned behavior/technology acceptance model, model of PC utilization, innovation diffusion theory, and social cognitive theory) [7].

A fundamental differentiation between variants of technology acceptance models and the approach proposed in this current research is the factor of time in which review of the information system being studied occurs. TAM [8], [9] and the extended Technology Acceptance Model (eTAM) [10] are tools which evaluate the perceived ease of use of technology application adoption (such as PDAs) however, the time when the process of managing an information system adoption occurs is made by reviewing prior actual adoptions, investigating variance of perceptions and applying the results to subsequent implementations. Liang et als study of usage of eTAM to predict actual PDA usage among healthcare professionals is such an example of this traditional review past implementations, apply findings to new/next implementation technology acceptance approach [11]. The approach taken in the ePOC project addresses the issues surrounding traditional approaches to user acceptance measurement by mapping concepts from agile systems development (DSDM) to user acceptance (UTAUT). This means that user acceptance is measured at each stage of development of the proposed system, enabling the project team to intervene at regular stages throughout the system development lifecycle to ensure that user requirements are being met and therefore increase the likelihood of user acceptance of the delivered system.

This approach is consistent with Toleman et als assertion that perception takes place through a sequence of stages through which a potential adopter of an innovation passes before accepting the innovation [12]. The ePOC project sought to gain end-user perceptions of a proposed new electronic point-of-care system, highlighting points where intervention by systems designers and project managers may be required to ensure the highest possible levels of user acceptance and support for the proposed system. In addition, more informal interaction was encouraged through a discussion forum on the project website and

notice board at the research site. Clinicians are actively encouraged to raise questions and/or concerns regarding the proposed system. Depending on the nature and imperative of the concern, the project team will either provide an immediate response or schedule it for discussion at a proceeding workshop. This enables intervention at critical points in the system development lifecycle.

The overall aim of this process for the ePOC project is to enhance user acceptance, the goal being to achieve an acceptance level (UA) as close as possible to 100%.

Methodology

User Surveys and Focus Groups are the predominant methods of data collection utilised in the ePOC project. This paper presents the results of five (5) surveys and the focus groups.

Development of the primary survey instrument

Klines Groupware Adoption Scale [13], [14] was adapted to meet the needs of the project. Klines original five subscales (EOU, Technical Support, Training, Work Needs Met, System Capabilities and Consultation), with a seventh subscale Commitment were included in the questionnaire. An eighth subscale, Persuasion was included in the questionnaire after the first survey.

The initial pre-implementation questionnaire was distributed to TACT five months (into the 3yr project) after project initiation, but well before the introduction of a prototype PDA device into the unit. The aim of the questionnaire was twofold; firstly to gauge the levels of perception by intended end-users towards to ePOC project and PDA device as the platform of choice and secondly, to validate the survey instrument. Results of the survey indicated significant numbers of TACT staff took a neutral stance on many questions (that is, responded Neither Agree nor Disagree). A focus group was conducted shortly after to attempt to understand the reasons behind this response. The outcome of the focus group revealed the assumption by the project team that TACT staff fully understood the aim of the ePOC project, their role in the project and that staff were receiving comprehensive updates on the project was not the case. Information passed on to the TACT Medical Director and Nurse Unit Manager at project committee meetings had not been conveyed to staff at TACT.

It was deemed imperative to the success of the ePOC project that staff at the unit be involved throughout the project and that management at the unit demonstrate a positive attitude and commitment to the proposed system. This outcome triggered a response by the project team to adopt a more concerted, consultative approach to system design, development and project management. An agile model of project management was adopted with the aim of increasing user acceptance throughout system development. This was facilitated through a series of surveys, focus groups, discussion sessions, newsletters, seminars and hands-on exercises with the PDA device scheduled to coincide with the release of each module (sub component) of the ePOC system. The approach taken is depicted in Figure 1.

An example of a module for ePOC is Clinical Observations. This module replicates current paper-based data entry capture of a patients clinical observations (such as Pulse, Respiration Rate, Blood Pressure and Blood Sugar Level, plus others) electronically on a PDA screen. The Clinical Observation module was developed by taking into account existing paper-based information requirements then user interface design considerations and workflow improvements. The prototype ePOC clinical Observation module was then tested in a field trial to determine if the module met expectations at point-of-care. TACT staff were surveyed and the results analysed to determine if, as a result of the project teams intervention and the new approach taken had enhanced end-user acceptance. Consider the Clinical Observation as Module 1 in Figure 1.

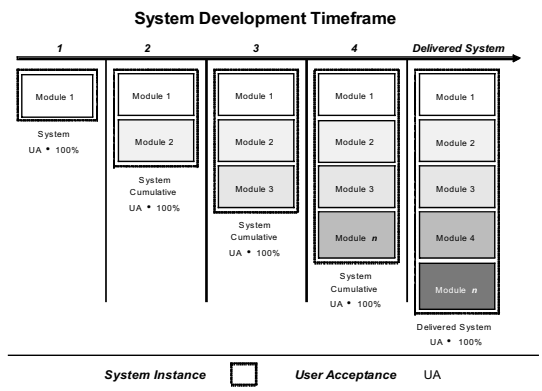


Figure 1 - ePOC Approach to System Acceptance

This process was repeated for subsequent modules; Peripheral Line Checklist, Electronic Protocols and MIMS on PDA. MIMS - an acronym for Monthly Index of Medical Specialties is an online prescribing guide for doctors, nurse prescribers and pharmacists. Significant findings with regards to each of the sub-scales of the survey instrument are discussed in the results section. Additionally, aggregated results of each of the five surveys and the sequencing of each in respect to points of intervention by the project team to clarify identified problems with each of the modules, impact on work practices and transfer of concepts onto the PDA device by TACT staff are shown in Table 1 on the following page.

Results

Results are categorised according to Klines subscales and the items included in the instrument. Aggregated results from the five surveys (questionnaires) are detailed in Table 1 and discussed below. Due to page limits imposed, discussion of results is limited to aggregated results across the survey subscales. Results of individual scale items are available on request. The first two surveys were conducted to assess user perceptions of the proposed Ambulatory Care Information System. Due to the fact that a large num-

Table 1- Aggregated Results (Positive Perception Levels)

	Ease of Use	Training	Technical Support	Consultation	Work Needs Met	System Capabilities	Commitment	Persuasion
Questionnaire 1								
Project & PDA	71%	84%	86%	76%	72%	62%	87%	N/A
Intervention 1								
Questionnaire 2								
Project & PDA	89%	98%	100%	93%	86%	86%	100%	87%
Field Trial 1								
Questionnaire 3								
Module 1	83%	54%	83%	50%	61%	71%	84%	67%
Intervention 2								
Questionnaire 4								
Module 2	94%	88%	100%	100%	89%	91%	100%	87%
Questionnaire 5								
Module 3	90%	91%	96%	90%	84%	88%	93%	90%

ber of TACT staff responded with a neutral response (neither agree/disagree) to all of the subscales in the first survey, a project champion (Medical Director, TACT) was appointed to drive the project at the client site. The purpose of this intervention was to determine if the appointment of a project champion would have a positive or negative effect on user perceptions of the proposed system. A seventh subscale, Persuasion was included in the questionnaire for the second survey to measure the impact of the intervention. The results of the second survey indicate that user perceptions across all subscales increased significantly, therefore demonstrating a positive outcome.

Following the second survey, TACT staff were provided with PDAs and given time to familiarise themselves with the devices prior to implementation of the first system modules, Clinical Observations and Peripheral Line Checklist. A field trial followed, after which a third survey and focus group were conducted. The results of the field trial indicate that staff acceptance of the system *decreased* after they had used it in the field. The focus group revealed a number of design issues related to the system and the user interface. The issues of concern raised by staff during the focus group were considered by the development team and the modules modified. Feedback from intervention 2 indicated that users were much more satisfied with the Clinical Observations module after their concerns over the design of the module and the user interface were addressed. Lessons learned form the previous survey and intervention were incorporated into the redesign of Module 2, Electronic Protocols (standardized work practices/ checklists based upon prescribed treatment of common diseases, for example steps to be undertaken in the treatment of Pneumonia or Cellulitis). Following implementation of the module onto the PDA device, staff again surveyed. The results of this survey record the highest levels of acceptance of all previous surveys providing further support for the approach to user acceptance taken by the project team.

The results of the fifth survey (Module 3, MIMS on PDA) also exhibit high levels of user acceptance across all 8 subscales, with increases in acceptance in the areas of training and persuasion. The other 6 subscales were down slightly on the previous survey. Notwithstanding, the results are promising. A focus group is planned for the near future to get feedback from users on the issues that emerged from the survey. This will enable further intervention and refinement of Module 3 before its final implementation on the PDA device. As a result of the feedback from staff at the unit, the project development team is currently working on graphical representation of the Clinical Observations and Peripheral Line Checklist modules. The surveys and focus groups will be repeated once the modules have been modified.

Conclusion

The results of the surveys and focus groups clearly demonstrate the viability of the approach to user acceptance taken in the ePOC project. Intervention by the project team at identified points in the development of the system has clearly resulted in increasing levels of user acceptance. Involving users in systems design has demonstrated benefits. Taking this concept a step further and applying it to user acceptance measurement throughout the system development lifecycle ensures that users will not only be more satisfied with the functionality of the system, but also its utility. Measuring user acceptance throughout the development of a system rather than post-implementation provides developers with an opportunity to address users concerns at the point where they arise, making it easier and more cost effective to modify system components. While research into user acceptance of the Ambulatory Care system is continuing, the results so far are more than pleasing. Future research will further develop the user acceptance framework and test and validate it in an enterprise-wide system development and implementation. It is conceivable that this approach can be applied to any system implementation that utilises a component-based system design and development methodology.

Acknowledgments

This research is derived from A PDA (Personal Digital Assistant) based Point of Care (POC) e-Health Solution for Ambulatory Care Services, an Australian Research Council Linkage Grant based at the Centre for Health Service Development, Wollongong University. The support of Pen Computer Systems, The Ambulatory Care Team (TACT) Northern Illawarra, South Eastern Sydney Illawarra Area Health Service (SESIAHS), Flinders University and the University of South Australia is acknowledged. Further information regarding the ePOC PDA project can be found on the ePOC projects website at the following URL:

http://www.itacs.uow.edu.au/cear/ehealth/ePOC/

References

[1] NSW Health Strategy for the Electronic Health Record. Report of the Health Information Management Implementation Coordination Group. February 2004. Available as PDF Online: <URL: http:// www.health.nsw.gov.au/iasd/chime/newspdf/ehr_doc.pdf > Last Accessed 11 October, 2006.

[2] NSW Health Community Health Information Management Enterprise: Overview: What is CHIME? Available Online: <URL http://www.chime.gov.au/iasd/chime/chime/ overview/index.html >Last accessed October 7, 2006.

[3] MedPDA.net, Medical PDA Reviews: Doctors using PDAs make less mistakes? October 6, 2004. Online <URL: http:// www.medpda.net/archives/001050.html > Last accessed September 19, 2006.

[4] DSDM Consortium. History. Available Online: <URL: http://www.dsdm.org/about/history.asp > Last Accessed 01 December, 2006.

[5] Venkatesh V, Morris, MG, Gordon B, and Davis F. User acceptance of information technology: toward a unified view. MIS Quarterly 2003: 27(3): 425-54.

[6] Dynamic Systems Development Methodology, Wikipedia. Available Online <URL: http://en.wikipedia.org/wiki/ Dynamic_Systems_Development_Method > Last Accessed November 2, 2006.

[7] Wade M and Schneberger S. Theories used in IS research: Unified Theory of Acceptance and Use of Technology. York University, Canada. Available Online: <URL: http://www. istheory.yorku.ca/UTAUT.htm > Last Accessed 03 November, 2006.

[8] Bagozzi R P, Davis F D and Warshaw P R. Development and test of a theory of technological learning and usage. Human Relations, 1992: 45(7): 660-686.

[9] Davis F D, Bagozzi R P and Warshaw P R. User acceptance of computer technology: A comparison of two theoretical models. Management Science, 1989: 35, 982-1003.

[10] Venkatesh V and Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Management Science. 2000: 46(2): 186-204.

[11] Liang H, Xue Y and Byrd T. PDA usage in healthcare professionals: testing an extended technology acceptance model. International Journal of Mobile Communications 2003: 1(4): 372389.

[12] Toleman M, Ally M and Darroch, F. Aligning Adoption Theory with Agile System Development Methodologies. PACIS2004, July 8-11, Shanghai, China. 2004. Available Online: <URL: http://www.sci.usq.edu.au/staff/markt/ papers/PACIS2004.pdf#search='Aligning%20 Adoption%20Theory%20with%20Agile%20System%20De velopment%20Methodologies' > Last accessed September 29, 2006.

[13] Kline TJB. The Groupware Adoption Scale: a measure of employee acceptance. Human Systems Management, 2001: 20: 59-62.

[14] Kline TJB and Gardiner H. The successful adoption of groupware: Perceptions of the users. Human Systems Management, 1997: 16: 301-306.

Address for correspondence

Dr Lois Burgess
Centre for Health Service Development
University of Wollongong
Northfields Ave,
2522, Wollongong, NSW
Australia

Email: lburgess@uow.edu.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

When Usage and User Satisfaction Differ: The Case of an Electronic Discharge Summary

Thomas Bürkle[a], Philip A. Engel[b]

[a] *Department of Medical Informatics, University Erlangen-Nürnberg, Germany*
[b] *ADMED GmbH, Cologne, Germany*

Abstract

We describe the results of a longitudinal study regarding system use and user satisfaction before and after introduction of an electronic discharge letter application in a pediatric intensive care unit (PICU) of a German university hospital. The new discharge letter application is part of the hospital information system (HIS). The study covered an eleven month time period and used system logs as well as questionnaires including a modified questionnaire of user interaction satisfaction QUIS. We used methods which are comparable to a previous study examining a HIS based discharge letter in three departments of an Austrian hospital. In comparison we found out that user satisfaction was lower in our case. Interestingly, we noticed that in our case this was mirrored by an increasing use of the new discharge letter although there was no pressure to switch to the new HIS based discharge letter application.

Keywords:

evaluation, clinical information system, discharge letter

Introduction

Evaluation of information and communication technology (ICT) which is used in a healthcare environment, is required to avoid potential side effects, because ICT can be inappropriately specified, have functional errors, be unreliable, user-unfriendly, ill-functioning or the environment may not be properly prepared to accommodate the ICT in the working processes [1]. The choice of evaluation methods and technologies is wide [2] and depends on available resources, goals of the evaluation and type of the technology to be examined [3]. ICT changes the ways how persons perform activities and influences the working environment over a long period of time. Within this study we were interested in usability and benefit of a ward specific discharge letter application within a hospital information system. Expectation of the users, extent of system use and user satisfaction were the parameters we measured before, during and after the introduction of the discharge letter application. Therefore we selected an interventional study design with three survey periods in order to display the development of the mentioned parameters over time. We used methods derived from other studies [4,5,6 and mainly 7,8] in order to compare our results with the results of other sites and researchers. Our study design tends towards a subjectivist approach [9], but we used classical observation methods only for a small part and relied mostly on questionnaires.

Materials and methods

The Münster University Hospital pediatric ICU

The pediatric intensive care unit PICU of Münster University Hospital has 16 beds and cares for all pediatric departments like pediatric oncology, pediatric cardiology and general pediatrics. A majority of the patients are premature babies and cardiosurgical cases. The PICU has 2.25 senior physician positions, 9.5 physician positions and 37 nursing positions. Nurses and physicians work in three shifts. In 2003 450 patients have been treated with a mean stay of 10.2 days per patient.

The hospital information system

Münster University Hospital introduced the commercial hospital information system Orbis® in 2000 [10], which is today sold by Agfa® and used in more than 400 German hospitals. Orbis® supports the development of user specific applications with a generator tool [11]. Technically Orbis® uses a document based approach. Electronic forms such as a lab result report, a radiology request or a discharge summary can be designed with the generator tool and may look similar to a paper based equivalent. At Münster University Hospital a variety of applications have been developed using the generator tool [12-14]. From our experience, the standard discharge letter which comes with Orbis® is not used in any department because it supplies a form which is almost empty and does not support the user with precompleted data from the patient record. Therefore we developed specific discharge letters for a variety of departments using the generator tool [11]. A ward specific discharge letter is based upon the typical discharge letter setup of the respective ward, it mirrors its structure and subdivision and supplies much prefilled patient information such as recent lab values, discharge diagnoses, last medication etc.

The study

A longitudinal single group non randomized intervention study using questionnaires and system logs was established with 3 survey periods. The intervention consisted of introducing a ward specific discharge letter application within the Orbis® HIS for the pediatric ICU in September 2003. The Orbis® system logged timestamps of first opening and last editing of each discharge letter. This data was analyzed in the study. The use of the new application was not mandatory, physicians could choose to use the old fashioned method of writing discharge letters in MS Word® on the same PC and to store them in a common directory. Discharge letters traditionally written in MS Word® couldn't be monitored reliably, as they were written based upon a previous letter and often subsequently stored under the old filename. Thus they were not considered for the study.

Three different types of questionnaires were distributed at three different points in time, namely before intervention in July 2003 (t1), five months after intervention in January 2004 (t2) and ten months after intervention in June 2004 (t3). Observation objects were all persons writing discharge letters, namely all physicians working regularly at the pediatric ICU in the respective questionnaire period as well as the ward secretary. While formal staffing counts 2.25 senior physician and 9.5 physician positions plus 1 secretary, there is considerable fluctuation due to physicians rotating among the various pediatric departments and some physicians hold part time positions only. We had a return of 14 questionnaires in the first round, 17 in January 2004 and 18 in June 2004. Questionnaires were marked with a numeric code to assure that answers of the same person could be traced through all 3 sampling periods. The follow up group which completed all 3 questionnaires had n=10 participants (9 physicians and 1 secretary)

The questionnaires comprised between 5 and 6 pages and were split into 4 chapters. Parts of the questionnaire were derived from previous studies on discharge letter applications [4,5,6,7,8]. In the first chapter we grouped questions regarding demographic data such as age, profession and time working at the PICU. The second chapter concerned expectations from (at t1) respectively experiences (at t2 and t3) with the electronic discharge letter. Statements were to be answered with yes or no. Besides it contained, similar to Ammenwerth and Kaiser [7,8], a part where the user could assess different activities like writing, modifying and searching discharge letters on a five point Likert scale. The third chapter, available only at t2 and t3, contained 17 questions of the modified and abbreviated QUIS questionnaire [5,6,7,8], which was also measured with a five point Likert scale. For the original QUIS refer to [17]. The fourth chapter contained questions regarding usability of other HIS functions. In the results' section of this paper we will only refer to the first three chapters. The questionnaire was each time complemented by open questions for free commenting. Any question could be skipped by marking "no answer" to avoid bias. Answers on the five point Likert scale have been converted to numerical values between 1 (do not agree at all) and 5 (do fully agree). All statistic results are descriptive only. We will also report some observational results which were made when the discharge letter workflow on the PICU was examined.

Results

Demographics

Our respondents in the follow up group (n=10) had a median age of 37 years and had been working 1-3 years (median) at the PICU at t2. At t3 the median stay time at the PICU had shifted to 4-10 years. One person had been working for over 10 years at the PICU. All ten had previous experience with computers. Of a total of 14 respondents at t1 twelve owned a PC themselves.

System use

Figure 1 shows the system use in terms of cumulated PICU discharge letters written in Orbis® during the study. A total of 568 discharge letters were recorded in the study period between August 13th, 2003 and June 29th, 2004 (48 weeks or 336 days). The corresponding graph shows a steady rate of approx. 7.04 discharge letters per week until January 04, afterwards we noted an increase to 15.3 letters per week.

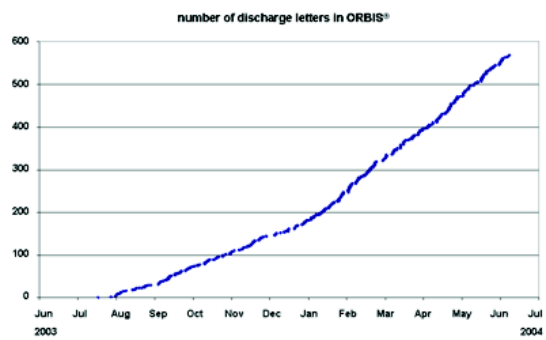


Figure 1 - system use: letters written with Orbis®

Expectations and experience

The expectations and experiences chapter was part of all three questionnaires. It comprised four questions to compare the expected benefits with those realized when the application was active (see also [4]). At t1 the question would e.g. read "I can imagine that the Orbis® discharge letter will be time saving for me" while at t2 and t3 we asked "Did you save time using the Orbis® discharge letter?". In a similar fashion the other three questions examined how easy it was to learn the use of the new application, if the application would make work easier and if it would improve the quality of patient care. All four questions had to be answered either with "yes", "no" or "Don't know". Results are given in table 1 for the follow up group with n=10 participants. The table uses an index of approval which is calculated as the relation between the number of positive answers and all answers. If all 10 participants would have answered positively, this index would

have been 1.0, if all users would have expressed a negative opinion it would have turned out 0.0.

Table 1 - Expectations and Experiences, index of approval

Index of approval	expected	experienced	
	t1	t2	t3
Easy learning	0.5	0.8	0.7
Easier work	0.9	0.6	0.7
Time saving	0.9	0.4	0.5
Better quality	0.7	0.0	0.1
Total	0.8	0.5	0.5

For time saving and easier working we measured high expectation values of 0.9 at t1, a slightly lower value for quality improvement and an indecisive (0.5) opinion regarding ease of learning. After system introduction at t2 and t3 the experienced values were found to be much lower (near 0.5) both for time saving and workload improvement while learning the use of the discharge letter application turned out to be easier than expected. At t3 experienced workload improvement scored a little higher at 0.7 but didn't reach the 0.9 at t1 which reflects the high expectation in the new application. Hardly any participant noticed an improvement in quality of patient care at t2 and t3.

Results of QUIS

Table 2 demonstrates the results gained with the modified QUIS [5,6,7,8]. QUIS was only included at t2 and t3, n is the total number of valid responses received for the respective question and varies for t2 (17 returned questionnaires) and t3 (18 returned questionnaires).

Table 2 - QUIS results at t2 and t3. Median, 25% Quantil and 75% Quantil and (for comparison) weighted mean values

No	Question	t2					t3				
		n	Q0.25	**Q0.5**	Q0.75	X	n	Q0.25	**Q0.5**	Q0.75	X
c11	Extra work with no visible benefit	16	4.0	**4.0**	4.3	3.71	18	2.3	**4.0**	4.0	3.33
c12	Easy to learn	17	3.0	**4.0**	4.0	3.47	18	3.3	**4.0**	4.0	3.67
c13	Others have benefit but not me	16	3.8	**4.0**	4.0	3.41	17	3.0	**3.0**	4.0	2.94
c14	Essential functionality is missing	16	2.0	**2.0**	3.0	2.18	18	3.0	**4.0**	4.8	3.72
c15	Well adaptable to needs	15	2.5	**3.0**	3.5	2.65	17	2.0	**3.0**	3.0	2.67
c16	Use is simple and self explaining	17	2.0	**3.0**	4.0	3.06	18	2.0	**3.0**	4.0	3.00
c17	Effort of use is appropriate	15	2.0	**2.0**	4.0	2.47	16	2.0	**2.5**	4.0	2.50
c18	Letters and results faster available	17	3.0	**4.0**	4.0	3.47	18	2.0	**3.5**	4.0	3.17
c19	Helpful persons to turn to available	16	2.0	**3.0**	4.0	3.00	17	2.0	**3.0**	3.0	2.56
c110	User training is unsatisfactorily	17	2.0	**3.0**	4.0	2.94	14	2.0	**3.0**	4.0	2.22
c111	My needs were respected	16	3.0	**3.0**	4.0	3.00	16	1.8	**3.0**	3.3	2.33
c112	User surface is nonuniform	15	2.0	**3.0**	4.0	2.76	14	2.3	**3.0**	4.0	2.44
c113	Makes information transfer easier	16	2.0	**3.5**	4.0	2.88	15	3.0	**4.0**	4.0	2.89
c114	Unreliable system	16	3.0	**4.0**	4.0	3.29	15	3.0	**4.0**	4.0	3.00
c115	Makes letter writing easier	12	2.0	**3.0**	4.0	2.00	13	2.0	**3.0**	3.0	2.00
c116	Is helpful for clinical research	11	1.0	**2.0**	4.0	1.53	14	2.0	**2.0**	2.0	1.56
c117	Few chances to do things wrongly	16	2.0	**3.0**	4.0	2.88	16	2.0	**2.0**	3.0	2.22
	Summary user interaction satisfaction					3.04					2.89

The table contains median, 0.25 Quantil and 0.75 Quantil as well as mean values (for later comparison with [8]). Values of six negatively phrased questions have been reversed to achieve a constant scale between 1 (do not agree at all) to 5 (do fully agree) for all QUIS questions. In order to compensate for questions which were not answered by all participants, mean values have been weighted according to [8] for the calculation of the summary user interaction satisfaction.

The results show a remarkable decrease in user satisfaction for several statements. E.g. statement c14 reads "There are essential functionalities missing in the discharge letter application". We note an agreement of 2="do not agree" for this statement at t2, five months after system usage. However at t3 after ten months system use, most users agree with this statement at 4="do agree". Question c117 reads "There are few things which can be made wrong when using the system". At t2 this statement was seen neutral at 3. Five months later users disagreed at 2="do not agree".

Accordingly the summary user interaction satisfaction score, calculated according to [8] on the base of weighted mean values, shows a decline from 3.04 at t2 to 2.89 at t3.

In table 3 we demonstrate the results concerning the assessment of different activities connected with the discharge letter (part of chapter 2 of the questionnaire). Results in a converted five point Likert scale are given for the follow up group with n=10.

Table 3 - median user satisfaction discharge letter activities

Activity	User satisfaction		
	t1	t2	t3
Write new letter	2.0	3.0	3.0
Modify letter	3.0	3.0	3.5
Sign letter	4.0	3.0	3.0
Search letter	2.5	4.0	4.0

Apart from the sign procedure for letters, results at t2 and t3 after introduction of the Orbis® discharge letter are equal or better in comparison to results that were received when writing discharge letters in MS Word®. Most values are slightly positive at or above the median value 3="don't know". Searching a discharge letter seems to have improved with values of 2.5 before and 4.0 afterwards.

Discussion

The results of this study indicate an increased usage of a HIS integrated discharge letter function within a pediatric intensive care unit although there was no pressure to switch from the previous method of writing discharge letters in MS Word®. Even at the end of the study, some very complicated discharge letters were still written in MS Word® but the increasing amount of Orbis® discharge letters clearly indicates that physicians decided to preferably use the HIS function. The total number of 568 discharge letters in 336 days written in Orbis® correlates fairly well with the average number of patients treated per year in this unit considering the fact that for most patients only one discharge letter is written. Furthermore, on request of the pediatricians the discharge letter function has in the meanwhile been spread to other pediatric wards successfully.

This increased usage is in contrast to a decrease in user satisfaction with the new electronic system measured with the QUIS part of the questionnaire. Expectations which had been set into the new application have not been met (see table 1). On the other hand user responses with regard to performing different activities such as writing, modifying and searching discharge letters seem to be more positive at t2 and t3 (values between 3 and 4 on a Likert scale).

The study is investigative rather than confirming, no hypotheses have been set or confirmed. The study design tends to be subjectivistic. From an objectivistic viewpoint there are some weak points in study design. It was non randomized and went over a long time period of 11 months. An influence of increasing workload and other factors cannot be excluded. We have a low number of study participants although we did include all staff concerned with the discharge letter application in the PICU. We enforced a high questionnaire return rate with the help of a senior PICU physician who assisted in distributing and collecting the questionnaires. This could have lead to bias in non responders or persons who are indecisive. We tried to minimize this effect by offering a no response opportunity for each question. Due to the small number of participants we did not do a formal evaluation with test-retest or split half methods for reliability and we do not have a formal gold standard to assess validity for the questionnaires. We may however state that the modified QUIS part (the original can be found in [17]) was evaluated and used in an identical fashion for the assessment of discharge letter functions by Ohmann, Boy and colleagues [5,6] and then again in a study by Ammenwerth and Kaiser [7,8]. The latter determined Cronbachs Alpha for reliability of this part with an excellent value of 0.9257 (max would be 1.0) and found good correlation of 0.78 between general satisfaction and QUIS determined satisfaction as an indicator for reasonable validity. Furthermore Ammenwerth and Kaiser [8] used similar questions for the assessment of activities performed with the new application (table 3). Other techniques (table 1) were used with good success in a previous study by ourselves [4]. While QUIS values were calculated on top of all respondents at t2 and t3 respectively, results given in table 1 and table 3 were calculated for the follow up group with n=10 only. However we did compare some of those questions for the follow up group and all residents at times t1, t2 and t3 exemplarily and found only small differences in the respective satisfaction values.

When looking at results of our study we may state in comparison to Ammenwerth and Kaiser [8] that our respondents were less happy with the discharge letter function. In our case summary user interaction satisfaction (weighted mean value) was between 3.04 at t2 and 2.89 at t3 whereas Ammenwerth and Kaiser measured higher values between 2.80 (neurology department) and 3.73 (transplantation surgery) among physicians in different units and 3.23 (transplantation surgery) as well as an extraordinary high value of 4.18 (neurology) among secretaries. In their publication [8] activities such as writing, modifying and searching discharge letters range among 1.67 (modifying letters in neurology) to 5.0 (secretaries searching letters in transplantation surgery) with an average of 2.83 for neurology and 4.17 in internal medicine. Our values from table 3 compare best to the neurology department which is least happy in the Ammenwerth and Kaiser study.

Ammenwerth and Kaiser explain differences between departments with organisational aspects in writing discharge letters. They assume that the department which previously had the best discharge letter workflow (neurology) reported the highest additional effort to switch to the new discharge letter and scored the lowest user interaction satisfaction. When examining the answers to open questions, we found comments that the layout in the Orbis® discharge letters was less favorable than in Word, that users complained about too much data (e.g. lab values) which was included automatically in the Orbis® discharge letter and that those letters were not specific enough. During observation we noticed that writing a PICU discharge letter at Münster university hospital is a fairly complicated

process where several physicians work on one letter during several days. In open questions some physicians complained that the long lasting procedure of assembling the discharge letter leads to overtime work which demonstrates a certain degree of general dissatisfaction with the process [19].

These observations could be compatible with a situation where a complicated workflow (namely writing a PICU discharge letter) was well accomplished with an established method (copying an old discharge letter of another patient in MS Word® and modifying it) and the introduction of a new method (writing discharge letters in Orbis®) was not perceived as an organizational improvement. But then, why did the physicians, without being forced to do so, switch to the new method for most cases and did even ask to spread the new method to more pediatric wards?

Conclusion

A subjectivistic study design like the one presented here will not confirm or deny any hypothesis but rather intends to focus on interesting aspects which may then be examined with objectivistic or further subjectivistic methods. In our case it revealed an interesting discrepancy between system use and user satisfaction. More studies will be required to examine differences in workflow before and after system introduction as well as other external influencing factors which can explain for the unexpected result.

Acknowledgements

We would like to thank all colleagues of the Münster University Hospital PICU, our special thanks go to Dr. I. Hörnig-Franz und Mrs. A. Kulms.

References

[1] Ammenwerth E, Brender J, Nykänen P, Prokosch HU, Rigby M, Talmon J. Visions and strategies to improve evaluation of health information systems - Reflections and lessons based on the HIS-EVAL workshop in Innsbruck. Int J Med Inf. 2004;73:479-491.

[2] Friedmann CP, Wyatt JC. Evaluation methods in Biomedical Informatics. Springer Science and Business Media New York 2006.

[3] Bürkle T, Ammenwerth E, Prokosch HU, Dudeck J. Evaluation of clinical Information Systems - What can be evaluated and what cannot? J Eval Clin Pract. 2001 Nov;7(4):373-385.

[4] Bürkle, T, Kuch R, Prokosch HU, Dudeck J. Stepwise Evaluation of Information Systems in an University Hospital. Methods Inf Med 1999;38:9-15.

[5] Ohmann, C, Boy O, Yang Q. A systematic approach to the assessment of user satisfaction with healthcare systems: Constructs, models and instruments. Stud Health Technol Inform. 1997;43 Pt B:781-785.

[6] Boy O, Ohmann C, Aust B, Eich HP, Koller M, Knode O, Nolte U. Systematic evaluation of user satisfaction of physicians with a hospital information system - initial results. Stud Health Technol Inform. 2000;77:518-22.

[7] Ammenwerth E, Kaiser F, Wilhelmy I, Hofer S. Evaluation of user acceptance of information systems in health care - the value of questionnaires. Stud Health Technol Inform. 2003;95:643-648.

[8] Ammenwerth E, Kaiser F. Evaluation der Benutzerzufriedenheit mit der rechnergestützten Arztbrief- und Befundschreibung. UMIT Schriftenreihe 1/2002, Innsbruck 2002.

[9] Moehr JR. Evaluation: salvation or nemesis of medical informatics? Comp Biol Med 2002;32:113-125.

[10] Prokosch HU, Bürkle T, Frankewitsch T, Ganslandt T, Gieshoidt K, Janßen U, Müller M, Lange M, Lüders D, Saile F, Vaclavik M, Veltmann U. Vom Krankenhauskommunikationssystem zur Elektronischen Krankenakte: Eine Statusbeschreibung zum KIS-Ausbau am Universitätsklinikum der Westfälischen-Wilhelms-Universität Münster. Informatik, Biometrie und Epidemiologie in Medizin und Biologie 2001;32:252.

[11] Kuhn KA, Lenz R, Elstner T, Siegele H, Moll R. Experiences with a generator tool for building clinical application modules. Methods Inf Med 2003;42(1):37-44.

[12] Bürkle T, Ganslandt T, Tübergen D, Menzel J, Kucharzik T, Neumann K, Schlüter S, Müller M, Veltmann U, Prokosch HU. Guideline Based Structured Documentation: The Final Goal? Stud Health Technol Inform 2002;90: 256-261.

[13] Müller ML, Bürkle T, Irps S, Roeder N, Prokosch HU. The diagnosis related groups enhanced electronic medical record. Int J Med Inf. 2003;70(2-3):221-228.

[14] Kunze U, Ganslandt T, Preusser P, Bürkle T, Prokosch HU, Senninger N, Haier J. Ein interdisziplinäres Modell zur onkologischen Dokumentation in einem Klinikinformationssystem. Zentralbl Chir. 2004;129(1):10-13.

[15] Chin J. Development of a tool measuring user satisfaction of the human-computer interface. In Chi'88 Conf. Proceedings: Human factors in computing. Association for Computing Machinery New York 1988:213-218.

[16] Ammenwerth, E, Gräber S, Herrmann G, Bürkle T, König J. Evaluation of health ISs-Problems and challenges. Int J Med Inf. 2003;71(2/3):125-135.

[17] Engel PA. EDV im Krankenhaus: Evaluation der Einführung der elektronischen Arztbriefschreibung am UKM. Inaugural Dissertation Universität Münster 2004. Available under http://miami.uni-muenster.de/servlets/DerivateServlet/Derivate-2177/01_diss_engel.pdf

Address for correspondence

Thomas Bürkle, Lehrstuhl für Medizinische Informatik, Krankenhausstrasse 12, 91054 Erlangen, Germany.
thomas.buerkle@informatik.med.uni-erlangen.de

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Mapping Clinicians' Perceptions about Computerized Protocol Use to an IT Implementation Framework

Shobha Phansalkar[a], Katherine A. Sward[b], Charlene R. Weir[a,d], Alan H. Morris[c]

[a] *Department of Biomedical Informatics, University of Utah, Salt Lake City, Utah, U.S.A*
[b] *College of Nursing, University of Utah, Salt Lake City, Utah, U.S.A*
[c] *Director of Research, LDS Hospital, Pulmonary Division, Salt Lake City, Utah, U.S.A*
[d] *Associate Director, Education and Evaluation, VA Salt Lake City GRECC*

Abstract

Previous studies have described the determinants of successful information technology (IT) implementation. In 2003, Kukafka et al. integrated several theoretical perspectives and proposed a framework for IT implementation. This framework is applicable to IT implementation in general but lacks the identification of factors affecting adoption, which are specific to the technology under consideration. We developed and validated a model that specifically identifies factors associated with clinicians' adoption of computerized protocols. The purpose of this paper is to identify the relations between the specific factors associated with intention to use computerized protocols and the high level variables that constitute the framework proposed by Kukafka et al. Incorporation of a specific model into a general schema for IT implementation allows implementers to assess the specific individual, organizational and environmental changes required to bring about successful implementation of computerized protocols. An understanding of clinicians' perceptions specific to the technology in use will allow its seamless integration into an organization's healthcare IT plan. Strategic planning requires enhancing the framework with additional detail related to the specific technology under consideration.

Keywords:

computerized protocols, clinician perceptions, implementation, clinical protocols [MeSH], decision making [MeSH], factor analysis [MeSH]

Introduction

Computerized protocols are knowledge-based systems used to provide decision-support to clinicians, for management of patients. Computerized protocol assisted decision support can range from receiving guidelines for disease management[1] to receiving individualized, detailed and explicit instructions, delivered at the patient's bedside, that are driven by dynamic patient-specific data[2]. Our use of the term 'computerized protocols', in this study, refers to explicit protocols that are capable of providing the latter, more specific support.

Explicit computerized protocols are capable of providing decision-support in a variety of complex clinical situations [3-6] at the bedside of the patient [2, 3]. Despite their advantages, resistance to adoption of evidence-based medicine, also labeled as 'clinical inertia', has impeded the implementation of protocol based decision support systems[4, 5]. Focusing on the technological aspects of information systems rather than their behavioral, organizational or attitudinal impacts has been identified as a cause for information system failure[6, 7]. To counter this resistance several behavioral models have been used for guiding implementation strategies[8]. Grimshaw and colleagues suggested the exploration of theory-based evaluations alongside randomized trials of dissemination and implementation strategies[9]. Such theoretical approaches not only provide a framework for understanding specific determinants of provider behavior but also allow exploration of causal mechanisms of adoption.

In 2003, Kukafka et al. proposed an IT implementation framework that integrated variables from prominent behavioral theories and models[10]. This multi-factor model can help in understanding the determinants involved in implementation and successful adoption of information systems in a health-care organization[10]. This framework is, however, applicable to IT implementation in general and previous studies have expressed the need for research on factors affecting adoption that are specific to the technology under consideration, so as to improve their predictive ability [11-13]. There is a paucity of literature that specifically identifies the cognitive and attitudinal factors associated with clinicians' adoption of computerized protocols [5, 14]

Previously, we developed a multivariate model of clinicians' behavioral intention to use computerized protocols; and we developed and validated a survey instrument, based on the model, that measures clinicians' behavioral intention to use computerized protocols [15]. This was a task-specific model focused on computerized protocol use, rather than information systems in general. The purpose of this paper is to identify the relationships between the specific factors associated with intention to use computerized protocols and the more general, high level framework proposed by Kukafka et al. Incorporation of the specific

model into a general schema for IT implementation places the specific model within a larger context and allows implementers to assess the specific individual, organizational and environmental changes required to bring about successful implementation of computerized protocols.

Methods

Development and validation of a model predicting clinicians' intention to use computerized protocols

A 'Grounded Theory' approach [16] was adopted for the development of a model predicting clinicians' intention to use computerized protocols [15]. This approach is used extensively in the social sciences and is based on an inductive rather than deductive approach to theory development. We conducted semi-structured interviews of clinicians with extensive experience in the use and development of explicit computerized protocols at LDS Hospital, in Salt Lake City, Utah. Five physicians, 3 nurses and 6 respiratory therapists were interviewed. The results of this work are published elsewhere[17] but we will summarize them briefly here. Three reviewers examined the transcripts looking for themes. After substantial group discussion, 39 themes were identified. The themes were then reduced and categorized into constructs derived from Expectancy-Value Theories[18], especially the Theory of Planned Behavior[19] and Intrinsic Motivation theories[20]. Inter-rater reliability for matching themes to constructs, was measured using Cohen's kappa ($k = 0.48$) and consensus was attained following discussion.

The constructs of the model for predicting computerized protocol use were subsequently used to develop items for a survey instrument. For each of the 8 constructs, we constructed up to 5 items. We administered the instrument to 240 physicians, nurses and respiratory therapists from University of Utah Hospital, Veterans Affairs Medical Center, Salt Lake City, and Intermountain Healthcare (LDS Hospital and Cottonwood Hospital). Factor analysis identified nine factors that accounted for 66% of the total variance cumulatively. Factors identified were: Beliefs regarding Self-Efficacy, Environmental Support, Role Relevance, Work Importance, Beliefs regarding Control, Attitude towards Information Quality, Social Pressure, Culture, and Behavioral Intention. The strongest predictor was Beliefs regarding Self-Efficacy, which accounted for 26% of the total variance of intention to use explicit computerized protocols. Results supported the reliability and construct validity of the instrument[15].

Mapping constructs predicting computerized protocol use to the IT implementation framework

The validated factors identified above were mapped to the IT implementation framework proposed by Kukafka et al. The IT implementation framework is adapted from Green and Kreuter's PRECEDE and PROCEED model [21]. Kukafka et al. proposed the adoption and application of the original model to IT implementation. This model identifies educational and organizational strategies for IT implementation. Underlying these strategies are three categories of factors, namely predisposing, reinforcing and enabling factors, which if modified, will be most likely to result in behavior change, i.e. adoption of the technology.

Results

The constructs from the model for predicting computerized protocol use [15] were mapped to the Phase 4 (Educational and Organizational) assessment in the IT Implementation Framework. Outlined below is the mapping of the specific constructs related to computerized protocol use to the Phase 4 assessment factors in the IT implementation framework. We have also included the definitions for the specific constructs.

- **Predisposing Factors**: These are factors related to the characteristics of an individual that motivates behavior change.
  - ***Work Importance***: An individual's tendency to orientate and value work in general.
  - ***Role Relevance***: The degree to which the computerized guidelines are relevant to one's perceived role at work.
  - ***Beliefs regarding Control***: The belief that that the behavior in question is in the control of the person.
  - ***Beliefs regarding Self-Efficacy***: The degree to which one believes that one has the skill to effectively engage in the behavior.
  - ***Attitudes towards Information Quality***: The attitude of the individual towards the quality of information that drives the logic of the computerized protocol.
- **Reinforcing Factors**: These are extrinsic factors in the form of rewards or punishments that are anticipated as a consequence of behavior change.
  - ***Social Pressure***: The perceived pressure that an individual might face from peers or supervisors that might facilitate or impede the behavior change.
- **Enabling Factors**: These are characteristics of the environment that facilitate the change by providing the skills or resources needed to bring about the behavior.
  - ***Culture***: This refers to the specific norms of the organization and whether the behavior change is in keeping with the goals of the organization itself.
  - ***Environmental Support***: The degree to which the environment is perceived as supportive to bring about the behavior change.

The specific relationships of the constructs to the overarching factors from the IT implementation framework are detailed in Figure 1. Figure 1 also illustrates how the specific model predicting behavior change fits into the realm of the general schema for IT implementation.

Discussion

Understanding user adoption of information systems is crucial for successful implementation. To facilitate this understanding several behavioral models have been employed for the adoption of information systems. A rich body of literature, incorporating models from various

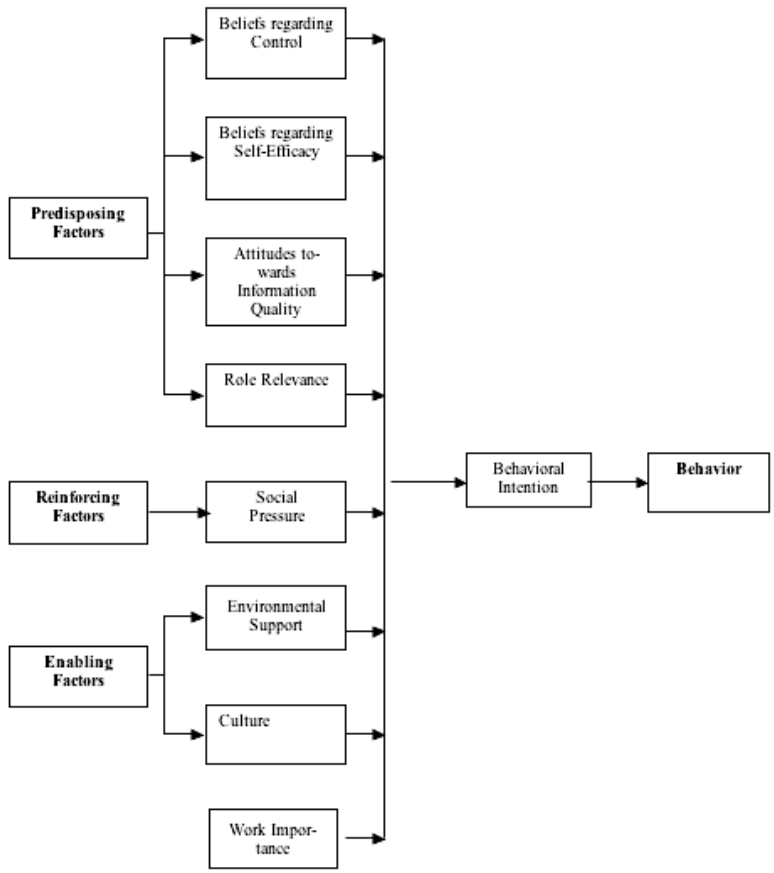


Figure 1 - Mapping of constructs related to computerized protocol use to the IT implementation framework

domains such as social and cognitive sciences, psychology, and the business sciences, has emerged. The constructs proposed in each of these models allow system implementers to understand the barriers to end-user adoption and design strategies for successful implementation of these systems.

Kukafka et al.[10] synthesized this literature and proposed an integrated multi-level framework for IT implementation. This framework is the first complex, multidimensional model that takes into account a variety of individual as well as organizational factors. Additionally this model identifies the factors that influence behaviors linked to IT use.

We developed and validated the model and the corresponding survey instrument specific to computerized protocol use. We then assessed how our model fits into the general schema for IT implementation, as proposed by Kukafka et al. Two constructs of Education and Policy Organization from the Kakafka model did not map onto any constructs in our model. For each of the predisposing, reinforcing and enabling factors, we identified specific constructs related to computerized protocol use.

The framework proposed by Kukafka et al. is designed as a foundation for driving IT implementation efforts in general. This IT implementation framework is designed to help IT implementers assess the individual and organizational changes needed for installment of a new innovation. However, identification of specific constructs related to the use of the technology under consideration need to be developed. These specific constructs allow IT implementers to make judgments about how the specific technology would fit into their healthcare IT plan. Also, identification of changes related to a specific technology would result in being able to make informed judgments about the timeline for implementation for a new technology. The IT framework serves as a guideline but needs explicit underlying constructs in order to make it readily usable. Our model for predicting computerized protocol use adds rich detail to the IT implementation framework, and can help explain, how specific predisposing, reinforcing, and enabling factors determine behavior related to use of computerized protocols.

This paper identifies how the IT implementation framework can be adapted for use with computerized protocol implementation. Identification of constructs related to

other technologies and how these can fit into the high-level factors proposed in the framework need to be assessed.

Conclusion

Incorporation of a specific model into a general schema for IT implementation allows implementers to assess the specific individual, organizational and environmental changes required to bring about successful implementation of computerized protocols. An understanding of clinicians' perceptions specific to the technology in use will allow its seamless integration into an organization's healthcare IT plan. Strategic planning requires enhancing the framework with additional detail related to the specific technology under consideration.

Acknowledgments

I acknowledge the contributions of Dr. Homer Warner for valuable help and discussions that led to the development of the model and this paper. Supported in part: NIH/ NHLBI ARDS Network: No1-HR-46062.

References

[1] A. Woodward, M. Wallymahmed, J. Wilding, G. Gill. Successful cardiovascular risk reduction in Type 2 diabetes by nurse-led care using an open clinical algorithm. Diabet Med 2006;23(7):780-787.

[2] A. H. Morris. Rational use of computerized protocols in the intensive care unit. Crit Care 2001;5(5):249-254.

[3] M. Vogelzang, F. Zijlstra, M. W. Nijsten. Design and implementation of GRIP: a computerized glucose control system at a surgical intensive care unit. BMC Med Inform Decis Mak 2005;5:38.

[4] R. Grol, J. Grimshaw. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003;362(9391):1225-1230.

[5] L. S. Phillips, W. T. Branch, C. B. Cook, J. P. Doyle, I. M. El-Kebbi, D. L. Gallina, C. D. Miller, D. C. Ziemer, C. S. Barnes. Clinical inertia. Ann Intern Med 2001;135(9):825-834.

[6] S. A. Davis, R. P. Bostrom. Training end users: an experimental investigation of the roles of the computer interface and training methods. . MIS Quarterly 1993;17:61-85.

[7] G. Gery. Granting three wishes through Performance-Centered Design. Communication of the ACM 1997(July).

[8] S. E. McLean, L. A. Jensen, D. G. Schroeder, N. R. Gibney, N. M. Skjodt. Improving adherence to a mechanical ventilation weaning protocol for critically ill adults: outcomes after an implementation program. Am J Crit Care 2006;15(3):299-309.

[9] J. M. Grimshaw, M. P. Eccles, A. E. Walker, R. E. Thomas. Changing physicians' behavior: what works and thoughts on getting more things to work. J Contin Educ Health Prof 2002;22(4):237-243.

[10] R. Kukafka, S. B. Johnson, A. Linfante, J. P. Allegrante. Grounding a new information technology implementation framework in behavioral science: a systematic analysis of the literature on IT use. J Biomed Inform 2003;36(3):218-227.

[11] T. Williams, C. May, F. Mair, M. Mort, L. Gask. Normative models of health technology assessment and the social production of evidence about telehealth care. Health Policy 2003;64(1):39-54.

[12] L. M. Hilz. The informatics nurse specialist as change agent. Application of innovation-diffusion theory. Comput Nurs 2000;18(6):272-278; quiz 279-281.

[13] V. Venkatesh, F. D. Davis. A theoretical extension of the technology acceptance model: Four longitudinal field studies. . Management Science 2000;46(2):186-204.

[14] F. D. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. . MIS Quarterly 1989;13(3):319-340.

[15] S. Phansalkar, C. R. Weir, A. H. Morris, H. R. Warner. Clinicians' perceptions about use of computerized protocols: A multicenter study. Int J Med Inform. 2007;Accepted, In Press.

[16] A. Strauss, J. Corbin. Basics of qualitative research: Grounded theory procedures and techniques. Newbury Park: Sage; 1990.

[17] S. Satsangi, C. R. Weir, A. H. Morris, W. H.R. Cognitive Evaluation of the Predictors of Use of Computerized Protocols by Clinicians. In: Proceedings of the American Medical Informatics Association; 2003; Washington D.C.: Hanley & Belfus; 2003.

[18] M. Fishbein. Attitude and the prediction of behaviour. New York: Wiley. ; 1967.

[19] M. Fishbein, I. Ajzen. Belief, attitude, intention, and behavior: An introduction to theory and research. Reading, MA: Addison-Wesley; 1975.

[20] E. L. Deci. Intrinsic Motivation. New York: Plenum; 1975.

[21] L. Green, M. Kreuter. Health Promotion Planning. 2nd edition ed. Mountain View: Mayfield Publishing Co.; 1991.

Address for correspondence

Shobha Phansalkar, RPh, MS
Department of Biomedical Informatics,
University of Utah School of Medicine
26 South 2000 East, Suite 5700 HSEB
Salt Lake City, UT 84112-5750, USA
E-mail: shobha.phansalkar@hsc.utah.edu
Tel: +1-801-582-1565 Extn: 2221, Fax: +1-801-584-5640

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

e-Health in Scotland: Setting a Baseline for Stakeholder Alignment

Sharon Levy, Anne Casey, Alison Wallis

Royal College of Nursing Dundee /London/ Lothian, UK

Abstract

Gaining knowledge of nurses' attitudes towards and understanding of various aspects of the Scottish e-Health programme is vital for needed 'stakeholder alignment'. This paper is focused on the Scottish results from a large on-line survey carried out in 2006, across the UK. Key findings, identified through analysis of both qualitative and quantitative data, are discussed. Results suggest that overall there is willingness and enthusiasm to engage and to see the modernisation of the NHS in Scotland underpinned by advanced and effective IT systems. It also shows that nurses have clear ideas about how information technology could help them and their patients in delivering safe care that enhances the visibility of the nursing contribution to care outcomes. It is argued that results help in setting a base-line from which to judge the success or otherwise of the needed e-Health clinical change management programme within clinical settings.

Keywords:

e-Health, nursing, Scotland, stakeholders alignment

Introduction

Moving away from a reactive, crisis-management, acute-orientated care towards anticipatory, preventative and continuous care which is integrated and supports faster, safer, more efficient, patient-centred services is the required direction of travel set for the National Health Service (NHS) in Scotland [1]. e-Health, the National Programme for the management of information and technology in the Scottish NHS was noted as a key driver to achieving such vision, in line with recommendation made in the Kerr review [2]. That report suggested that the Scottish Executive refocus its e-Health strategy and appoint a visible and high profile leadership so as to ensure clinical buy-in. The Scottish Executive appointed, in late 2006, a full time e-Health Nursing, Midwifery and Allied Health Professions (NMAHP) lead to promote and support the needs of the largest group of professional care providers in the health service.

One of the key priorities for the NMAHP lead is to achieve what the Auditor General for Scotland has called 'stakeholders alignment'. This phrase is used in the published national audit report on management of IT in the Scottish health service [3], and is amongst the 10 'leading practices' set to support effective implementation of clinical IT systems. The report also notes that the Scottish Executive is taking steps to promote its e-Health strategy, "in line with good practice", and yet there is no signposting to available data that set a baseline from which to judge the success or otherwise of the needed stakeholders alignment. This paper aimed to fill this apparent gap, through reporting on the analysis of the Scottish data extracted from the largest UK nursing on-line IT survey undertaken to date [4]. It is argued that the presented results should be used to identify a development plan as well as to ascertain current levels of knowledge and engagement of the nursing professions with the e-Health national programme in Scotland. Adopting the findings as an agreed base line will enable the programme to chart future progress of their alignment with the nurses, health visitors and midwives who provide care for the population of Scotland.

Method

The method chosen to conduct our survey was an on-line format and we used the Nursix.com portal, which provided immediate quantitative data analysis and feedback. The questionnaire which was piloted and used in early 2004 [5] and developed further in 2005 [6] included 7 demographic questions, 23 multiple choice and 1 open ended question for free text input. The 2006 study commenced on 31st of May and lasted for 34 days. Participants were asked to take part in the study through direct e-mail approach (n = 50,000), promotion on the public part of the Royal College of Nursing (RCN) web site, the RCN on-line discussion zone and RCN Bulletin (circulated to all 392,000 members). The results from analysis of completed questionnaires (n = 4453) makes this study the largest nursing e-Health survey in the UK to date. Amongst the total UK sample were 415 Scottish nurses (9%) and this paper is focused on analysis of their response.

Despite the fact that the proportion of Scottish nurses within the sample population is similar to their proportion within the RCN membership, we are not claiming the sample is representative of the nursing, midwifery and health-visiting workforce in Scotland. This was a convenience sample of self-selected RCN members who are, nevertheless, more likely to be interested in and aware of NHS e-Health developments. After all, subjects willingly completed the survey which relied on a degree of competence in using a computer and the internet. This makes the reported findings about these nurses' level of e-Health awareness even more significant.

Results

In response to the question '*How much information have you had about NHS IT developments?*' 35% of respondents from Scotland felt they had fully or reasonably adequate information. This is a slight increase of awareness amongst subjects compared with 20% in 2004 and 29% in 2005. Nevertheless, 38% had inadequate information and a further 28% stated they had no information about these developments at all. A similar picture is noted with awareness of integrated electronic health record developments. Most (72%) nurses had inadequate or no information about such development at all. Interestingly, 42% of respondents said that e-Health developments were a 'Very Important' or 'Important' priority at their place of work with a further 31% who were 'unsure' about the corporate priority that e-Health gets within NHS Scotland. Nevertheless, one in three (32%) of the Scottish nurses who responded to the survey felt that spending large amount of money in the pursuit of e-Health was a poor (24%) or very poor use of NHS resources.

When asked how important is consultation with individual practicing clinicians about new NHS IT developments in Scotland, the majority of nurses believe that consultation was very important (55%) or fairly important (30%). However, when asked what consultation they have had about integrated electronic health care records, 69% told us that their views were not sought. It is worth noting that only 1% of respondents explicitly indicated that they do not wish to be consulted.

Despite many having inadequate information about proposed e-Health developments, many respondents (56%) believe that integrated electronic health care records will improve clinical care with 49% agreeing that using electronic patient health records will lead to significant (31%) or slight improvements to their nursing practice. Thirty percent of the nurses were 'unsure' about the effect and impact of such developments on their practice.

The final questions regarding integrated electronic health care records were set to identify possible effects on the confidentiality of care records and perceived impact on relationship with patients. Thirty nine per cent of respondents believe that there would be little or no effect on confidentiality of care records. However, a significant minority (31%) believe that the use of electronic records may pose a threat to confidentiality. Although 46% of these Scottish nurses felt that electronic records will have no or little effect on the therapeutic relationship with patients, 38% suggested that such development will have a beneficial impact.

Training and support needs of nursing staff who are expected to use new Scottish e-Health systems and products were also sought. As was noted in results from the 2004 and 2005 surveys, the vast majority (95%) of respondents indicated that training is 'Very Important' (83%) or 'Fairly Important' (7%) to the success of the Scottish e-Health programme. When asked how much training they had received in working time in the last six months, 71% had no training at all in working time (compared with 68% in 2005 and 69% in 2004). The majority (68%) of Scottish subjects also felt that around the clock technical support is essential to the success of the integrated electronic health care records with a further 20% indicating it was important or fairly important (6%).

Access to hardware was also included in the survey: respondents were asked how many people share a computer in their immediate clinical area. Interestingly, Scotland had the worst ratios for hardware with 15% needing to share access to a computer with more that 30 other colleagues with a further 9% sharing it with more than 20 people and 11% with more than 10 others.

Apart from gathering quantitative data, the online survey also gave participants an opportunity to add free text regarding 'anything else they wish to note'. A quarter of those who completed the questionnaire opted to add qualitative information and thus provided a personal reflection on different aspects of e-Health as it relates to their practice area and experience. The main themes to emerge from this qualitative data were **communications, access, training and support**.

Many nurses felt that they have too little information on current developments, with those who care for clients in the community developing a sense of abandonment: "Working as Community Palliative Care nurses, employed by a voluntary organisation but working as sole practitioners in a geographical area, we are not routinely copied into NHS information and miss out on new developments…NHS Trusts do not remember we are there!"

Community staff also seem to be unable to access training: "Staff within the community setting are all basically self learning…. We require more training and more computers and those we have do not allow us to contribute to patient records which would be very beneficial to the patient, nurse and doctor. I feel there is not enough emphasis placed on community involvement."

However, even those who work in acute settings report difficulties in accessing training: "Training for some nurses with no IT skills is essential if this is to be implemented efficiently and effectively. Many managers view IT skills for other departments rather than nursing..." "Not enough protected time to be trained, not enough access to computers, our one computer [situated in the ward office] is a work tool and as such is always busy, there is always somebody needing access to it, to check emails, send emails, make documents, compile records/files.." "In my clinical area, there is no computer access. Access is only available for more senior managers."

There are those who are fearful that nurses are driving down the 'high tech low touch' route: "I feel the Health service is becoming so impersonal, we are heading for supermarket health care and that is so sad, because we as nurses have so much more to give." Yet there are others who argue that nurses must embrace the 21st century: "Many people believe that nursing will suffer, but realistically a lot of the repetitive work carried out by nursing staff should be minimised, for example, how many times does a patient get asked their name and date of birth ???"

The great majority of nurses believe that e-Health will provide better and safer care: "The introduction of electronic patient records will ensure a more seamless and holistic approach to patient care particularly in the out of hours periods when often problem arise. The more information that is available the better and quicker the outcome of care can be made."

Being a very practical profession, nurses are aware of problems as well as possible ways to address the challenges ahead: "Systems need to be shared between Health and Social work and other partners who have access to providing any input to the patient, or they are no use. There will still be duplication of information and extra work if they are not". The striking finding is that many want to be involved, informed and engaged in e-Health developments so that both patients and clinicians are able to benefit from safe and efficient systems that support optimal care.

Discussion

Delivering 21st century care that is safe, effective and evidence based, relies in part on efficient use of clinical tools that deliver reliable, accurate, and timely information. Information and communications technologies (ICT) have the potential, in combination with organisational modernisation, to revolutionise the way care is provided. The Scottish e-Health strategy [7] is focusing on delivery of a common ICT system that will one day replace paper records. The strategy is also promising to connect every clinician to a secure health information network. There is a commitment to enabling healthcare professionals to access best practice guidance and knowledge at any point of care. However, the provision of modern healthcare in Scotland will require not only new infrastructure but also new thinking about practice and new skills to maximise the use of the clinical information superhighway.

Results from our on-line survey suggest that nurses in Scotland are not yet adequately prepared for the e-Health vision and thus unable to utilise or maximise the opportunities that ICT offers the profession. Despite the fact that the majority of Scottish nurses who responded to the survey see the potential of e-Health for both service users and care providers, they are as yet unable to integrate the use of ICT into their routine practice. Nurses do not get the right amount of training, nor are their views sought when systems are being designed, built, tested and implemented. The risks to the e-Health programme are grave. Poorly designed systems that do not fit with the requirements of the clinicians and with clinical workflow lead to inadequate implementations that can contribute to unsafe practice. Moreover, excluding the rich and holistic elements of nursing gives rise to task orientated products that restrict core element of professional care and affect the care outcome.

Closer collaboration between the professional bodies and the Scottish executive is urgently needed to provide a range of solutions to the needs identified by nurses in this survey. These include:

- Provide e-Health resources including adequate Information, advice and guidance
- Provide e-Health educational resources including a range of learning 'products', events and Continuous Professional Development (CPD) opportunities for e-Health
- Devise an educational policy that embeds e-Health competencies in both undergraduate and post graduate courses
- Resolve e-Health workforce issues so that the contribution nurses have on care outcome becomes visible
- Facilitate e-Health evaluation, research and development from the perspective of the nursing profession
- Grow and maintain e-Health capacity and leadership in nursing.
- Moreover, Scottish nursing must be an integral part of a UK initiative to develop an infrastructure and process for getting consensus on, maintaining and getting conformance on record content standards: the tools, charts, record structures etc that reflect best, evidence based nursing practice. The achievement of appropriate and agreed nursing content in Scottish electronic health record will ensure that the people of Scotland get the care from nurses that they expect and deserve.

Conclusion

As attitudes to computers were found to be a cardinal factor for effective use of IT in the work place [7], understanding nurses' attitudes towards many aspects of the Scottish e-Health programme is vital for planning effective change management within the clinical environment. This survey captured the views of a large number of Scottish nurses (n=415), albeit self selected. More research could be undertaken to see whether these findings are supported using a wider, more representative sample of nursing professionals. However, this paper clearly demonstrates that many nurses are keen to be involved in developments and implementation of all new information and record systems that have nursing components to demonstrate the unique value nursing brings to the clinical encounter.

Acknowledgments

The authors wish to thank those RCN members who were willing to share their views and aspiration with them.

References

[1] Scottish Executive Health Department, Delivering for Health, Edinburgh, Scottish Executive 2005.
[2] Scottish Executive Health Department, Building a Health Service Fit for the Future Edinburgh, Scottish Executive 2005
[3] Audit Scotland Informed to care: managing IT to deliver information in the NHS in Scotland, Edinburgh Audit Scotland 2006
[4] Royal College of Nursing, IT Survey 2006: Nurses and NHS IT developments, London, Royal College of Nursing June 2006 available at http://www.rcn.org.uk/publications/pdf/nurses_and_NHS_IT_developments_survey_2006.pdf

[5] Royal College of Nursing, Nurses and NHS IT developments: results of an online survey by Nursix.com, London, Royal College of Nursing; February 2004 http://www.rcn.org.uk/downloads/research/nurses-it-devs-survey.doc
[6] Royal College of Nursing, A year on: nurses and NHS IT developments London, July 2005; http://www.rcn.org.uk/downloads/press/nurses-it-devs-survey2005.doc
[7] National eHealth/IMT strategy 2004-8
[8] Jayasuriya, R. and Caputi, P. Computer attitude and computer anxiety in nursing: Validation of an instrument using an Australian sample. *Computers in Nursing*. 1996; 14(6): pp.340-345.

Address for correspondence

Sharon Levy, EPICentre, 9 Bell Street, Dundee, DD1 1HG Scotland.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Usability of Institutional Cancer Web Sites: an Italian Case Study

M. Cristina Mazzoleni[a], Raffaella Butera[a,b], Franco Corbella[c], Vittoria Balcet[c], Enrico Masenga[b]

[a] *IRCCS Fondazione Maugeri, Pavia, Italy*
[b] *ProgettareWeb.it, Italy*
[c] *IRCCS Policlinico San Matteo, Pavia, Italy*

Abstract

In order to evaluate if and to what extent Italian speaking cancer patients can benefit from information available on cancer web sites, an "in vitro" usability (ISO definition) study has been carried out. It investigated the usability of the web sites of the most representative Italian Institutions in the oncological field for the adult patients needing to find information about head and neck cancer. Specific evaluation criteria from the literature were used. The results point out some problems about accessibility, in line with other studies, and about the usefulness of the contents, in particular in the web sites of care delivery institutions: a grey present situation, but there are already grounds for significant improvement. Institutions and organizations must not waste the opportunity of being valuable sources in order to build the so called "informed patient," and the usability of their web sites could make the difference.

Keywords:

internet, medical informatics, usability, patient education.

Introduction

A more informed patient is unanimously recognized as a determinant for the improvement of both the population health status and the effectiveness of the healthcare systems. Patient information/education still represents a challenge, and the web is one of the emerging tools to achieve this goal. Initially, the major concerns dealt with the *reliability* of health information available on the web for the general public, but the discussion is still open [1]. In the last fifteen years many empirical studies, the great majority dealing with Anglophone web sites, have been carried out in order to evaluate the *quality* of information for consumers. In a patient/consumer-centered perspective *usability* is the next step and could/should be the keyword and the focus, provided that the content is "good."

The document *ISO 9241-11 (1998) Guidance on Usability* by the International Organization for Standardization defines usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". This definition implies the *usefulness* concept (*achieve specified goals with effectiveness*), while another authoritative figure - Jacob Nielsen - considers usability as a part of usefulness. The two constructs are anyway evidently strictly bound. Both of them refer to the potential benefit that the users can get from the evaluated object, in terms of result - to meet a need - and necessary effort to achieve it.

The present paper deals with an "in vitro" usability (ISO definition) study having for *object* the web sites of the most representative Italian Institutions in the oncological field; for *users* the patients, adults; for *need* to find information about head and neck cancer; for *context* a standard PC with the most popular operating system (Windows XP), browser (Internet Explorer 6.0), screen resolution (1240x768 and 800x600), with no optional plug-in. The study was aimed at evaluating if and to what extent Italian speaking cancer patients can really benefit from cancer web sites for a more active participation to the care process.

Materials and methods

Sample construction

To identify the most representative healthcare institutions in the oncological field, the list of the members of a national network of excellence named "Alleanza contro il Cancro" was used. To this first group some National Patient Associations and the Italian Association of Oncology Physicians were added because of their popularity and relevance. The choice of restricting the sample of web sites only to the ones published by authoritative well known institutions had the implicit assumption that the provided information was reliable.

Evaluation criteria

Evaluation criteria were based on the "Research-based Web Design & Usability Guidelines" [2] developed by the U.S. Department of Health and Human Services. These guidelines are primarily addressed to web site managers and are aimed at "creating better and more usable health and human services web sites". The approach taken to produce these guidelines is compliant with the traditional model used to build guidelines in the health care domain, supporting each statement with scores (range 1-5) related to both relative importance (relative importance score, RIS) and strength of evidence (strength of evidence score, SES). All of the items with RIS equal to 5 were taken into consideration, independently from the SES value (Table

1). The items were numbered consecutively (ID); the reference to the original guidelines items are listed in the second column of the table.

Table 1 - Items with Relative Importance Score equal to 5

ID	GL #	Guideline item heading		SES
1	1:1	Provide useful content		5
-	1:2	Establish user requirements		4
-	1:3	Understand and meet user's expectations		3
-	1:4	Involve users in establishing user requirements		3
2	2:1	Do not display unsolicited windows or graphics		3
3	3:1	Comply with section 508	Comply with W3C WCAG 1.0	2
	3:2	Design forms for users using assistive technology		2
4	3:3	Do not use colour alone to convey information		4
5	5:1	Enable access to the homepage		3
6	5:2	Show all major options on the homepage		2
7	5:3	Create a positive first impression of your site		4
8	6:1	Avoid cluttered displays		3
9	6:2	Place important items consistently		4
10	6:3	Place important items at top center		4
11	8:1	Eliminate horizontal scrolling		4
12	9:1	Use clear category labels		4
13	10:1	Use meaningful link labels		4
14	13:1	Distinguish required and optional data entry fields		3
15	13:2	Label pushbuttons clearly		2
16	15:1	Make action sequences clear		4
17	16:1	Organize information clearly		4
18	16:2	Facilitate scanning		4
19	16:3	Ensure that necessary information is displayed		2
20	17:1	Ensure usable search results		3
21	17:2	Design search engines to search the entire site		3

The items 1:2, 1:3, 1:4, addressing specifically the development process and were excluded. The items 3:1 and 3:2 were transformed into a single item "Comply with the World Wide Web (W3C) Consortium Accessibility Initiative Guideline (WCAG) 1.0" keeping the same meaning. The compliance was tested using an automatic evaluator [3]: web sites were considered compliant if W3C WCAG 1.0 Priority 1 level checkpoints were satisfied. For all of the mentioned items, two observers marked each web site as compliant or not compliant. In case of discordant judgments, a common result was obtained through a further joint inspection.

In order to evaluate web sites compliance with the item 1:1 ("Provide useful content", the only one with both RIS and SES equal to 5) usefulness was defined as the capacity of meeting the patient's information need. The results of a review [4] about information needs of cancer patients were used. In the quoted review the needs are divided into 10 categories and in 64 subcategories, after having analyzed the results of 91 articles published between 1980 and 2003. From the original list proposed by the authors of the review, the 7 categories quoted in at least 25% of the papers were extracted and considered in our analysis as primary patient's needs: disease-specific, treatment, prognosis, rehabilitation, coping, impact on interpersonal/social relations, and consequences on body image and sexuality. The 41 subcategories with a frequency of occurrence within the category grater than 5% (see the Results section for the complete list) were analyzed in detail. The presence of information about the selected subcategories and specifically related to head and neck cancers was checked by the two observers. For each web site, a score was attributed to each subcategory (subcategory score, SS) in case of presence of specific information (SS=1), presence in general but not related to head and neck cancers (SS=0.5), or absence of information (SS=0), respectively. These results were summarized at category level: each category was given a category score (CS) equal to 1 in case of presence of information (SS greater than 0) in more than 50% of the subcategories, 0 otherwise.

Results

The sample

Fourteen web sites were identified. One of them required optional software for displaying Java applets in order to start navigation: being out of the test constraints, this site

was excluded, due to its total inaccessibility in a standard situation. The final sample was composed of eight web sites managed by research and care delivery institutions (RCI) and five by patients/physicians/research associations (PPRA). The list of evaluated web sites is available form the authors on request.

Compliance with items 2 to 21

Table 2 shows the percentages of web sites compliant with the usability requisites listed in Table 1, with item 1 ("Provide useful content") excluded. As to item 14, no data entry was required in any of the tested pages: no evaluation, consequently, was possible. Search functionality (items 20 and 21) was present in 11 out of 13 web sites: the percentages shown are referred only to the reduced sample. Search results were very poorly displayed (item 20), with only one exception in which the retrieved information was provided in a very effective way. In general evaluated web sites provided only basic functionality. As to the accessibility requirements, 38% (5 out of 13) of the home pages were sealed as not compliant with the W3C WCAG 1.0 Priority 1 level. Within this group of non accessible sites, three belong to PPRA.

Table 2 - Web sites compliance with usability requisites 2-21

Requisite	% (*)
2. Do not display unsolicited windows or graphics	93
3. Comply with the W3C WCAG 1.0	61
4. Do not use color alone to convey information	100
5. Enable access to the homepage	93
6. Show all major options on the homepage	100
7. Create a positive first impression of your site	100
8. Avoid cluttered displays	100
9. Place important items consistently	100
10. Place important items at top center	93
11. Eliminate horizontal scrolling	100
12. Use clear category labels	100
13. Use meaningful link labels	100
14. Distinguish required and optional data entry fields	-
15. Label pushbuttons clearly	100
16. Make action sequences clear	100
17. Organize information clearly	100
18. Facilitate scanning	93
19. Ensure that necessary information is displayed	86
20. Ensure usable search results	16
21. Design search engines to search the entire site	100

* percentage of compliant web sites

Compliance with item 1 "Provide useful content"

Table 3 shows the distribution of subcategory scores (SS) for RCI and PPRA web sites. Limiting the analysis to information specifically related to head and neck cancer (SS=1), only 10 subcategories out of 41 were present in more than 6 web sites.

Figure 1 shows the comparison between the content available on RCI and PPRA web sites. The bars represent, for each category, the percentage of web sites in which at least 50% of the subcategories were present.

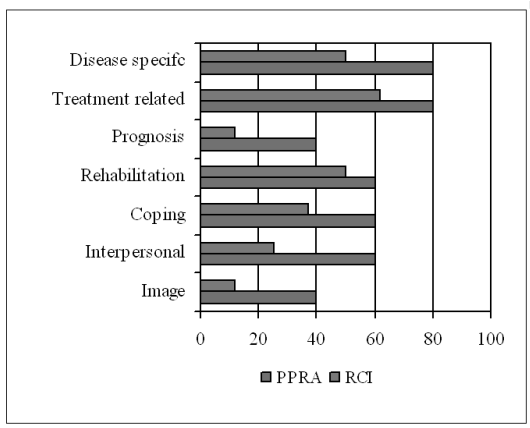


Figure 1 - Percentages of PPRA and RCI web sites with CS=1

Discussion

There are several methods to analyze the usability of a software, and hence of web sites. Some tools have been developed and validated to assess the subjective perception of the users [5,6], but not many experiments have been carried out to evaluate usability according to well defined criteria and guidelines. The "Research-based Web Design & Usability Guidelines" proved to be a good tool to inspect the usability of health related web sites, joining technical and content-related aspects. The choice of considering only the requisites with RIS equal to 5 has led to the exclusion of other items such as those related to readability and comprehensibility of the texts (RIS equal to 4). These requisites, when dealing with health related web sites, are in our opinion even more critical than others quoted in Table 1. A readability test is ongoing for a sample of retrieved texts. Fairly good results are expected, in accordance with a previous study performed in 2003 [7], in

which 75% of the analyzed texts were easily readable for Italian people who have attended compulsory education.

Overall, all of the web sites inspected met the great majority of the usability basic requisites: this could be due to the great simplicity of most of them in terms of structure and provided functionality. However, it should not be forgotten that one web site was excluded from the sample due to the impossibility to display it in the standard PC configuration defined for our study. As regards accessibility, although automatic evaluation is quite a rough method, it is suitable for identifying at least inaccessible pages. The results indicate a lack of attention paid to web accessibility, despite a recent law for the public administration web sites and many recommendations published in the grey literature. Although referring only to a small sample, the situation depicted in this study is comparable with that reported in other studies [8,9,10] where the percentages of non accessible sites were respectively 60%, 66% and 65%.

As to requisite 1 "Provide useful content", the assessment took into account only the breadth of topics' coverage because of the a-priori assumption that the information came from authoritative institutions, and hence was reliable and accurate, if present, by definition. The fact that an institutional web site has necessarily to deal with all of the topics can be questionable. A patient develops, during the care process, trust in and familiarity with a particular institution: in our opinion he/she could expect the web site of this institution to be comprehensive and to be an effective tool to easily find the needed information. This doesn't mean that a single web site has to provide, by itself, all of the needed information, but it should provide a way to reach it. For this reason in the present study a topic was considered as being present both in the case of direct publication and in the case of delivery via a link to a specialized external web site.

The results point out that, globally, web sites provide more frequently information related to the medical culture (category DS, Disease specific; category TR, Treatment; category R, Rehabilitation). At the same time, the least present subcategories were the ones related to the prognosis (P), a topic that is ranked quite high in the patient's needs list.

Table 3 - Cancer patient's information needs and number of sites, out of thirteen, with subcategory score (SS) 0, 0.5 and 1

Category	Subcategories	SS=0			SS=0.5			SS=1		
		rci	**ppra**	**tot**	**rci**	**ppra**	**tot**	**rci**	**ppra**	**tot**
Disease-specific (DS)	Type of cancer/nature of disease	4	1	**5**	0	0	**0**	4	4	**8**
	Aetiology and course of disease	4	1	**5**	0	0	**0**	4	4	**8**
	Physical effects of disease	4	1	**5**	0	0	**0**	4	4	**8**
	Specific diagnosis information	4	1	**5**	0	0	**0**	4	4	**8**
	Stage of disease	4	1	**5**	0	0	**0**	4	4	**8**
	Symptoms of cancer/management of symptoms	4	1	**5**	0	0	**0**	4	4	**8**
	Where to get information about specific diagnosis	4	2	**6**	0	0	**0**	4	3	**7**
Treatment-related (TR)	Side effects, risks and benefits of treatment	3	2	**5**	2	1	**3**	3	2	**5**
	Available treatments/treatment options	4	1	**5**	1	1	**2**	3	3	**6**
	Treatment plan/description/logistical info	4	1	**5**	0	0	**2**	3	3	**6**
	Tests and procedures involved in treatment	4	1	**5**	1	1	**2**	3	3	**6**
	Reducing side effects of treatment	5	4	**9**	1	1	**2**	2	0	**2**
	Alternative or complementary treatments	5	2	**7**	1	1	**2**	2	2	**4**
Prognosis (P)	Chance of cure	6	2	**8**	0	0	**0**	2	3	**5**
	Life span or survival rate	7	2	**9**	0	0	**0**	1	3	**4**
	Recurrence of cancer	7	2	**9**	0	0	**0**	1	3	**4**
	Spread of disease or metastasis	7	3	**10**	0	0	**0**	1	2	**3**
	Expectations for future health condition	8	4	**12**	0	0	**0**	0	1	**1**
	Effect on life plan or long term goals	8	4	**12**	0	0	**0**	0	1	**1**
	Outcome of no treatment or delayed treatment	8	5	**13**	0	0	**0**	0	0	**0**

Category	Subcategories	SS=0			SS=0.5			SS=1		
		rci	ppra	tot	rci	ppra	tot	rci	ppra	tot
Rehabilitation (R)	Self care issues or home care during recovery	5	3	**8**	0	0	**0**	3	2	**5**
	Nutrition during recovery	5	3	**8**	0	0	**0**	3	2	**5**
	Immediate post-treatment follow-up care	5	4	**9**	0	0	**0**	3	1	**4**
	Long-term side effects of cancer or treatment	6	4	**10**	0	0	**0**	2	1	**3**
	Recognizing/preventing treatment complications	5	4	**9**	0	0	**0**	3	1	**4**
	Recovery time	5	3	**8**	0	0	**0**	3	2	**5**
	Where to get medical supplies/equipment	4	3	**7**	0	0	**0**	4	2	**6**
	Maintaining physical health or physical activity	5	2	**7**	0	0	**0**	3	3	**6**
	Prevention and early detection	2	0	**2**	0	0	**0**	6	5	**11**
	Maintaining psychological health	2	0	**2**	0	0	**0**	6	5	**11**
	Health behavior and promotion	2	0	**2**	0	0	**0**	6	5	**11**
Coping (C)	Emotional reactions/support, coping	5	2	**7**	1	0	**1**	2	3	**5**
	Community counseling or support	5	2	**7**	1	0	**1**	2	3	**5**
	Support groups	6	2	**8**	1	0	**1**	1	3	**4**
	Support from other patients	6	2	**8**	1	0	**1**	1	3	**4**
Interpersonal / social (IS)	Effect on family, friends, or caregivers	5	2	**7**	1	0	**1**	2	3	**5**
	Effect on social life or leisure	5	2	**7**	0	0	**0**	3	3	**6**
	Risk of disease for family members	8	4	**12**	0	0	**0**	0	1	**1**
	Effect on employment or work life	6	2	**8**	0	0	**0**	2	3	**5**
Image / sexuality (I)	Sexuality	7	3	**10**	0	0	**0**	1	2	**3**
	Physical appearance/physical attractiveness	7	3	**10**	0	0	**0**	1	2	**3**

The limited presence of this kind of information could reflect the caution usually used in Italy to deal with such problems outside the face to face patient-physician relationship. Cultural differences may play an important role [11] in determining the content of health related web sites. Moreover, within category P, one subcategory is completely absent: no information at all is provided as to the "outcome of no treatment or delayed treatment". The difference between RCI and PPRA web sites as regards content coverage is quite evident from the graph of Figure 1. Among the RCI web sites, three out of eight gave no information or minimal information about a very limited number of topics, neither directly or indirectly through links to other web sites. This is the reason why in Figure 1 the percentages of RCI web sites are never higher than 62%. The highest percentages were reached in those cases (categories Disease specific, Treatment, Rehabilitation) in which at least one RCI web site provides information through a link to the web site of a PPRA. Among the PPRA web sites, only one out of five did not provide health related information. The comparison with the results of other studies aimed at inspecting the completeness of health related information on web sites is not easy, due to differences in the used methods, in the sample building procedure, and in specific disease. In many papers the *quality* of the content has been evaluated comparing the information provided with the content of "golden standards" like published guidelines and recommendations. In the present study we made an attempt to evaluate the *usefulness* of the content, starting from the patient's information needs. In a study [11] some percentages are reported as to the presence of treatment related topics in samples (any kind of web sites, not institutional only, gathered through search engines) of English and German web sites: information about the most frequent therapy options (category TR in the needs schema) are present in 90% and 69% of the English and German samples respectively, compared to a figure of 69% (9/13) of the PPRA plus RCI total sample in our study. In [12] a description of US Children's Hospitals web sites is provided: the conclusion referred to the contents "users *[....]* would be disappointed by most of the sites" indicates a failure of institutional web sites in meeting the users information needs. In another paper [13] about Norwegian hospital web sites, health related contents are not even quoted, fact from which the absence of useful content, as defined in our study, could be derived.

Independently from the numbers, a certain disappointment as to the usefulness of the contents in the Italian Cancer Institutions web sites was unavoidable. A gleam of hope appeared when, surfing MEDLINE, a paper titled "More information more choice: an Italian database for oncology

patients" [14] was retrieved. It describes a "library for the patients" that was developed and put on line within a nation wide project named Azalea. Partners of this project were most of the institutions whose web sites have been inspected in the present study. The service, since August 2006, is no longer available due to reasons independent from its quality. It has been available for more than 2 years, with an increasing number of accesses (up to 20.000 unique visitors per month in the last trimester of presence on line). More than 2000 documents were available, dealing with most of the topics listed in Table 3, and integrated into the database after expert validation. Last but not least, a reassuring seal of WACG compliance was present in the homepage. The hope is that this valuable effort will not have been wasted.

Conclusions

When dealing with patients looking for information on the web, the usability of Institutional web site should be taken into great account, since, as stated by the Internet Healthcare Coalition, people are expected to use web sites managed by institutions or organizations in which they have confidence. The results reported in this paper point out some important usability problems in particular about accessibility as to patient associations web sites, and about the provision of useful content, as to the web sites of care delivery institutions. The present situation of the Italian Cancer Institutions web sites is grey, but luckily there are already the grounds, and maybe even more, for a significant improvement. Institutions and organizations must not waste the opportunity to be valuable sources in order to build the so called "informed patient," and the usability of their web sites could make the difference.

Acknowledgments

The authors thank Dr. Ivana Truccolo and Dr. Katia Bianchet of the Biblioteca Scientifica - Centro di Riferimento Oncologico Aviano IRCCS, and Dr. Ing Marco Pagani of CBIM, Pavia for having provided the data about the Azalea project.

References

[1] Adams SA, DeBont AA. Notions of reliability: considering the importance of difference in guiding patients to health care Web sites. Methods Inf Med. 2003;42(4):307-12.
[2] www.usability.gov/pdfs/guidelines_book.pdf
[3] www.webxact.watchfire.com
[4] Rutten LJ, et al. Information needs and sources of information among cancer patients: a systematic review of research (1980-2003). Patient Educ Couns. 2005;57(3): 250-61.
[5] Nahm ES, Resnick B, Mills ME. Development and pilot-testing of the perceived health Web site usability questionnature (PHWSUQ) for older adults. Stud Health TechnolInform. 2006;122:39-43.
[6] Kirakowski J, Corbett M: SUMI: the software measurement inventory. Brit J Educ Technol. 1993;24:210-212.
[7] Mazzoleni MC, et al. Italian speaking health consumers and the offer of information on the internet. Proceedings of "Cancer on the net", New York, 2003
[8] O'Grady L. Accessibility compliance rates of consumer-oriented Canadian health care Web sites. Med Inform Internet Med. 2005;30(4):287-95.
[9] Chiang MF, Starren J. Evaluation of consumer health website accessibility by users with sensory and physical disabilities. Medinfo. 2004;11(Pt 2):1128-32.
[10] Mancini C, Zedda M, Barbaro A. Health information in Italian public health websites: moving from inaccessibility to accessibility. Health Info Libr J. 2005;22(4):276-85.
[11] Weissenberger C, et al. Breast cancer: patient information needs reflected in English and German web sites. Br J Cancer. 2004;91(8):1482-7.
[12] Kind T, et al. Do the leading children's hospitals have quality web sites? A description of children's hospital web sites. J Med Internet Res. 2004;6(2):e20.
[13] Norem J, Moen A. The websites of Norwegian hospitals: do they meet national guidelines and patient's expectations? J Telemed Telecare. 2004; 10(5):272-6.
[14] Bogliolo A, et al. More information, more choice: an Italian database for oncology patients. Ann Oncol. 2005;16(12):1962-7.

Address for correspondence

M. Cristina Mazzoleni - Fondazione Salvatore Maugeri, IRCCS
Via Ferrata 8,
27100 Pavia,
Italy
email: cmazzoleni@fsm.it

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Investigating Internet Use by Mental Health Service Users: Interview Study

John Powell[a], Aileen Clarke[a]

[a] *Health Sciences Research Institute, Warwick Medical School, University of Warwick, United Kingdom*

Abstract

The internet is an increasingly important source of mental health-related information, and has the potential to be harnessed as a tool to support self-care and informed decision-making. Yet little is known about the motivations and attitudes of users. We therefore undertook a qualitative interview study with a purposive sample of mental health service users with internet experience, to explore issues with respect to mental health-related internet use. One of the prime motivations for online mental health seekers was to find experiential information from other people with similar problems. This information allowed users to know they were not alone, and to instill hope that others in the same situation had recovered. Benefits of the internet as an information source included convenience, privacy and anonymity. Problems related more to misuse of the internet rather than concerns over inaccuracy. Such qualitative work is important in an emerging research area to understand internet use better.

Keywords:

internet, mental health, qualitative research

Introduction

The internet is playing an increasingly significant role in health information-seeking [1]. At the same time health systems in developed countries are witnessing increasing consumer involvement in healthcare, with the recognition of the importance of health literacy, the developing role of the informed expert patient, and the increasing importance of self-management in chronic disease [2]. These developments can be facilitated and enhanced using e-health tools. An important first step to achieve this is to understand the motivations and attitudes of e-health users.

A recent UK study has shown that over 10% of the general population has used the internet for mental health information [3]. Online mental health interventions such as internet-based cognitive behavioural therapy are increasingly being developed [4]. In a qualitative study to investigate mental health information needs (reported elsewhere [5]) we asked additional questions related to mental health-related internet use to explore the issues in using this increasingly important medium. In particular we investigated the advantages and disadvantages of the internet as a source of mental health information, and explored user motivations and concerns.

Materials and methods

Recruitment

In this qualitative study, we undertook a series of in-depth interviews with adult mental health service users. As investigating internet use was one of our key aims we deliberately over-sampled individuals who had direct experience of using the internet for mental health information. Participants were identified and recruited purposively through various means. In primary care and secondary care settings in Oxfordshire, UK, recruitment used information sheets displayed in healthcare settings and given to potential participants by general practitioners and psychiatric staff. In addition, study advertisements were sent to mental health user organizations, and posted on a consumer health information website. We included participants aged 18 or over, with personal experience of mental health problems, and who had recent (or current) experience of mental health services. Participants were excluded if they were unable to take part in a one hour interview due to ill health.

Procedure

One interviewer (JP) conducted all interviews in person or via the telephone. A topic guide was used based on a review of the literature. Written consent was given and all interviews were audio-recorded and transcribed. Open-ended questions and follow-up prompts relating to internet use were used to explore issues related to online mental health information seeking. Questions inquired about benefits and problems of internet use for mental health information, and motivations for online information seeking. National Health Service ethics committee approval was given for the study.

Analysis

A grounded approach was used to identify themes and subthemes through a process of familiarization with the data, open coding, axial coding, and selective coding [6]. Two investigators (JP, AC) undertook the analysis. Themes related to general mental health information seeking have been described elsewhere [5]; this current paper presents the findings specifically with respect to mental health-related internet use.

Results

Participants

Thirty-six participants were purposively recruited and all consented to be interviewed. There were 25 females and 11 males from different points in the mental health system representing a variety of occupational backgrounds and with an age range from 25 to 64 years. We intentionally over-sampled individuals with some experience of using the internet (32 of 36 interviewees). This was not a representative sample, but a purposive one designed to explore specific issues.

Themes

Benefits of the internet

The first theme concerned the advantages of the internet as a source of mental health information. These can be summarised as anonymity, privacy, convenience, accessibility and empowerment.

Interviewees reported valuing the "unobtrusive" nature of the internet - both in terms of the anonymity it affords users and also the way that you can participate without interacting.

As interviewee 22 (a 38 year old social scientist) put it

> "It's such an unobtrusive, discrete way of doing it [finding information]. I mean you can do it very privately without even, having to talk to another person."

Interviewee 30 (a 52 year old building contractor) also described the advantages of the internet:

> "I think another advantage is that you can do it in the privacy of your own home because if you are conscious of the stigma or have difficulties with speaking about what you have with other people then, you know you can do it completely in your own privacy."

Interviewee 36 (a 34 year old shop manager) described the convenience (not having to go out), the possibility of interacting with other people in a similar situation (who may not be available in "real life"), and the fact that one can "eavesdrop" and not conform to the normal social rules of "real life" - for example by just leaving:

> "[Its] easy to find information, so much easier than say… I don't use a library, or books. You don't have to go out. And other people, I think, the sort of support from other people, there is always someone who has had your problem, whatever it is you know. Which is amazing. Its nice to know that you are not alone, the silly little things you see coming on line and saying, thank God I found this great pal, I'm not alone anymore, you know, what they thought was some obscure condition that no one else had because maybe they are the only person in their city that has got it or something. There is always loads of other people on the internet. And also you can just leave incidentally if you want to which you can't do in a real situation. … You can just sit there and eavesdrop (laugh) … but you couldn't really do that in a real situation, you can't just stand at a corner of the room and listen, people won't let you."

Interviewee 15 (a 30 year old freelance journalist) contrasts her experience of seeking help from her general practitioner and with that of internet help-seeking.

> "I tend not to be very assertive in things to do with myself and not, you know, not want to take up GP's time so I find it very difficult to ask for information that focuses on myself, whereas if I sit at the internet, you know, I am anonymous, nobody would, uhh, I am not taking up anyone's time and its just a lot easier, a lot easier and I can do it when I am ready to do it and I don't have to wait and get stressed about it. Also I don't have to worry about, if I ask a question someone is going to say 'well, why are you asking' you know, whatever and I don't get the third degree. So that's why I prefer to do it [on the] internet rather than in person."

There also seem to be particular benefits of the anonymous nature of the internet for people with mental health problems, as interviewee 32 (a 48 year old retail manager) explains.

> "From my point of view its [using the internet] because I cannot interact in a social group. That's my worst problem, one of my worst problems is being able to interact with a group of people. I find it absolutely impossible anymore and I hide myself away, I am becoming a hermit I would say. But I do go, I do go out, I have to go out to the doctors, I have to take my wife to the shop, I don't go in the shop, but she goes, but, you know, so its anonymity, I haven't said that right but you know what I mean, because they can't see me, if they could see me, I would stay away."

The interviews also demonstrated the benefit of the internet in improving access to information. Two elements were seen as beneficial: access to other people with the same condition was greatly facilitated by the global nature of the internet, and secondly, access to expert knowledge was valued and seen as empowering. For example interviewee 1 (a 58 year old civil servant) explained:

> "If patients can actually get information off the internet, they've got some, umm, something to argue with the doctor about, I think in terms of empowering people, being without information is disempowering and this evens up the power between the doctor and the patient and the doctors may think you are a nuisance simply because you arguing with me and there is an element of the doctor knowing best, but umm, I mean, that it, that is my view and certainly gathering information in the last few days felt quite armed to go in and sort out the psychiatrist simply because we got this information"

Problems of the internet

Regarding any particular problems with using the internet as a source of healthcare information, it was interesting that the accuracy of online information was less of a concern for interviewees than misuse of the internet. People recognised that there are websites with poor or inaccurate

information but this was not seen as a major issue. Interviewees acknowledged the unregulated nature of the internet and that some of the information was “silicon snake oil” (interviewee 11, a 43 year old mature student), but they seemed to have confidence in their ability to discern accurate sources and not be taken in by the “quacks” or “cranks”. Individuals had developed their own strategies for dealing with inaccurate information, and had learned to trust certain websites - usually those with identities that they trusted in the "real" world - such as sites run by the NHS, the BBC or major mental health charities. Misuse with disruptive or malicious intent was seen as more problematic. Interviewee 14 (a 43 year old care worker) commented on her fears of who might be using chatrooms for victims of abuse:

> "I have come across [websites] where they have got like chat rooms for people who have been abused and things like that and mental health stuff and I wouldn’t go in there because I am thinking 'ooh', you know, I don’t know whether it would be full of genuine people or whether it would be full of, you know, if there is going to be people in there that want to hear about people who have been abused and get a kick out of it."

Interviewee 16 (a 33 year old care worker) and interviewee 32 (a 48 year old retail manager) describe the disruption of websites they were using:

> "Then, over a period of time, the site seemed to be taken over by some very young teenagers particularly, some at boarding schools and everything, and they were using the chat rooms and the message boards making threats saying, 'I am going to kill myself now' and then they would sign off and things like that, and to me that wasn’t helpful … I know the internet is all about free speech and free opinion and everything but for somebody, yeah, I mean, I won’t deny, I, I have made three serious attempts on my life and a number of others and, I, I spent nearly three months in a coma and I have got liver damage and I live with that permanently and to then find sites on telling you how to commit suicide, I do find it somewhat distasteful (pause) and you, you had that on NetDoctor [a health information website] people were asking the best way to kill themselves and I can't deal with that, and a lot of people couldn’t deal with that."

> "Unfortunately you do get people who come on there that are not ill. You can tell they are not 'cause they start arguing, not arguing but leaving nasty comments and like I have just said there was one person not long [ago] who come on and say, 'you are not depressed, you are just this, that and the other, get your self together'. …When people come on and are being nasty like that, I don’t like that."

The major concern about using the internet for health information for our interviewees therefore concerned disruptive online behaviour, rather than poor quality information. A further minor complaint from our UK respondents concerned the predominance of US websites found when searching for mental health information, and the fact that the material on these sites was often not suitable for a UK context.

Hearing about other people’s experience online

The importance of hearing about other people’s experience of mental health problems, and using the internet to find these was the prominent motivation emerging from the interviews. In particular the internet met the needs of users to know that they were not alone with their problems (characterised as “universality”) and to know that others in a similar situation have been able to get better (characterised as “hope”). Interviewees also valued how the interactivity of the internet allowed them to obtain understanding and empathy from others in their situation.

Knowing one is not alone was seen as reassuring, and also it helped the individual reject their notion of 'madness'. As interviewee 2 (a 30-year old biochemist) put it:

> "I think it has helped in that, I understand that I am not the only person to act and behave in the way I do, Certainly since joining [name of website] its just so nice to read what people write, thank goodness I am not as mad as I thought I was."

In our interviews madness was seen as something individual, idiosyncratic, an extreme deviance from normality. "Some kind of alien" as interviewee 15 put it. Whereas if you have something that other people also have, then you no longer see yourself as 'mad'. As illustrated by interviewee 31 (a 55 year old occupational therapist):

> "I think the most helpful thing was to know that some other people had similar feelings and I wasn’t going mad."

The benefit of not only knowing that one is not alone, but also that other people have had the same problem and got better is illustrated by interviewee 13 (a 27 year old teacher):

> "As I got a bit braver and kind of went back to work and so on, I did the odd search on the internet and I would read other people's experiences of schizophrenia and it was often by their families, or sometimes it was by the person themselves, and just hearing that somebody had recovered would mean so much to me because when I was still kind of in the recovering process, I was heavily depressed thinking that I was never going to be the same person that I had been before, I became ill, and reading experiences where people had recovered, it was such a boost, because you thought well, if they can do it, you know, I will be damned if I can’t."

Not only did individuals want to know that they were not alone and that others had got better, but they also wanted to interact with others, because only other people who have been through the same experiences as them know "what it's like". Interviewee 34 (a 25 year old unemployed man) who had depression talked about his experience of using an internet bulletin board, and the value of interacting with others who also had this diagnosis:

> "They can understand and know what you are going through as well. … If you have not been through the experience you don't understand what it is."

Interviewee 5 (a 37 year old teacher) described how finding people who knew what she was going through on the internet had been a "lifeline":

> "It has been a lifeline literally… It makes a difference between being lonely and afraid and unable to do anything and being able to turn on a metal box and at the other end of it, there's people that know exactly what you are going through and they can support you through it because they have been through it and they have come through the other side or they are going through it still."

Finally, interviewee 16 (a 33 year old care worker) described how contact with peers in the same situation on the internet allowed her to fill an information gap that her doctor was unable to meet:

> "You can reach the end of your doctor's knowledge and then you can go online, and you can talk to other people who have also been treated for years and years and years and they can help you to come up with new ideas … For me, personal information is the most useful, you know, what it was like to take this particular drug, or what its like to have a particular condition because the person who has got it or who is taking that drug they can describe what it is really like and sometimes it then kind of makes sense. You may have seen that in the official description but you didn't really understand exactly what it meant."

Other motivations

Aside from finding experiential information, the other main motivations concerned undertaking personal research into the condition. Prime areas for research were the causes of illness, alternative diagnoses, and treatment options. Internet research (often in association with searching other sources) was frequently undertaken in response to a lack of information available from the health service. Interviewee 11 (a 43 year old mature student) explained:

> "Nobody told me anything, I have researched it myself. I knew a fair amount and I got that information from work, but, nobody offered me any information, nobody pointed me in the right direction I should say … I came across something in a book ... That's Borna disease virus. I didn't know whether this was anything relevant … so I then did some research through the internet. I ran a search on the name of the disease."

The following quote from interviewee 14 (a 43 year old care worker illustrates how this personal research links to the reported benefit of empowerment described earlier:

> "I diagnosed myself with fibromyalgia, my doctors hadn't done it and I did take stuff from the internet and leaflets and said, I think that's what I have got, you have not listened to me all these years, you have just said 'oh yes, we think you are depressed you know' … and this doctor actually read it and my symptoms and said 'oh you are very clever … you have diagnosed yourself' … and this was just information … We went to see the specialist and he said to me, the doctor has diagnosed.. and I said 'no I did it' and he said 'how' and I said 'I got information off the internet, looked up symptoms that I was suffering and then went to a library and got some information and then a girl by pure chance had a leaflet' … and apparently, I have just recently found out that depression is part of fibromyalgia."

Discussion

Studies in areas other than mental health have identified the benefits of the internet for health consumers of anonymity and convenience of access [7;8]. Our interview analysis supports these previous findings particularly around the advantages of "privacy". Privacy in this case encompasses both anonymity and the private access that people have to the internet in their own homes and it is important for the avoidance of stigma. The internet offers advantages in acting as a medium of mass communication whilst allowing for individual interaction. It provides both a public and a private space for information-seeking from a variety of sources while allowing the concealment of individual identity.

Our finding that users trust certain websites and that these tend to be related to organisations they would trust in the real world is also supported by other work [9;10]. A focus group study in eight European countries found that participants often reported a feeling of being overwhelmed by the volume of information on the internet, and had concerns about information quality [11]. While we found that individuals did express reservations about the reliability of online information, they were actually more concerned with internet misuse than with untrustworthy information. This is supported by a US population survey which showed that 81% of internet users expect to find reliable information about health or medical conditions online [12]. Indeed 46% of internet users in this survey said they would use the internet as the first source next time they needed reliable medical information, compared with 47% who would contact a medical professional. This is interesting in the context of the overwhelming volume of articles published in the biomedical literature expressing concern at the quality of online information.

Previous work has identified that patients exchange information with each other online [13;14], and there are emerging findings of the value of hearing other people's experience on the internet. In two qualitative studies re-analysing illness narratives collected for the DIPEx project, Ziebland and colleagues have shown that cancer patients use the internet (amongst other reasons) to find experiential information from other patients [7;9]. Hardey has discussed how his studies of interviews with internet users and analysis of internet-based illness narratives show that individuals use the internet both for finding out and displaying personal experiences as well as for professional information and advice, and how the sharing of experiences is part of a wider shift in the relationship between lay and medical expertise [15].

Conclusion

The internet is a valued source of information on mental health issues and users describe its benefits, particularly concerning privacy. It is meeting the need for a private space to discuss mental health issues. There are concerns about the internet; related more to misuse than to inaccuracy. In particular the internet is a source of information about other people's experience of illness, providing universality and hope. Healthcare providers wishing to harness the internet as a mental health resource need to take account of the motivations of users and their perceptions of the risks and benefits of mental health-related internet use.

Acknowledgments

Dr Powell was funded by a National Health Service Research and Development Programme Training Fellowship in Health Services Research.

References

[1] Baker L, Wagner TH, Singer S, Bundorf MK. Use of the internet and e-mail for health care information: results from a national survey. *JAMA* 2003;289:2400-6.
[2] Gray JAM. The Resourceful Patient. Oxford: eRosetta Press, 2002.
[3] Powell J, Clarke A. Internet information-seeking in mental health: population survey. *Br J Psychiatry* 2006;189:273-7.
[4] Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. *BMJ* 2004;328:265.
[5] Powell J, Clarke A. Information in mental health: qualitative study of mental health service users. *Health Expect* 2006;9:359-65.
[6] Strauss A, Corbin J. Basics of Qualitative Research. Techniques and procedures for developing grounded theory. 2nd ed. London: Sage, 1998.
[7] Ziebland S, Chapple A, Dumelow C, Evans J, Prinjha S, Rozmovits L. How the internet affects patients' experience of cancer: a qualitative study. *BMJ* 2004;328:564.
[8] Umefjord G, Petersson G, Hamberg K. Reasons for consulting a doctor on the Internet: Web survey of users of an Ask the Doctor service. *J Med Internet Res* 2003;5:e26.
[9] Rozmovits L , Ziebland S. What do patients with prostate or breast cancer want from an Internet site? A qualitative study of information needs. *Patient Educ Couns* 2004;53:57-64.
[10] Dutta-Bergman M. Trusted online sources of health information: differences in demographics, health beliefs, and health-information orientation. *J Med Internet Res* 2003;5:e21.
[11] Coulter A, Magee H. The European Patient of the Future. Maidenhead: Open University Press, 2003.
[12] Horrigan JB, Rainie L. The Pew Internet & American Life Project. Counting on the Internet. Washington, DC: The Pew Internet & American Life Project, 2002.
[13] Feenberg AL, Licht JM, Kane KP, Moran K, Smith RA. The Online Patient Meeting. *J Neurol Sci* 1996;139:129-31.
[14] Hoch DB, Norris D, Lester JE, Marcus AD. Information Exchange in an Epilepsy Forum on the World Wide Web. *Seizure* 1999;8:30-4.
[15] Hardey M. E-Health: The internet and transformation of patients into consumers and producers of health knowledge. *Information, Communication and Society* 2002;4:388-405.

Address for correspondence

Dr John Powell,
Associate Clinical Professor, Health Sciences Research Institute, Warwick Medical School, University of Warwick,
Gibbet Hill Road, Coventry CV4 7AL, United Kingdom.
Telephone: +44(0)2476 574883.
Email: john.powell@warwick.ac.uk.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Text Characteristics of Clinical Reports and Their Implications for the Readability of Personal Health Records

Qing Zeng-Treitler[a], Hyeoneui Kim[a], Sergey Goryachev[a], Alla Keselman[b] Laura Slaughter[c] Catherine Arnott Smith[d]

[a]*Decision Systems Group ,Brigham and Women's Hospital, Harvard Medical School, U.S.A.*
[b]*Lister Hill National Center for Biomedical Communications, National Library of Medicine, U.S.A.*
[c]*Center for Shared Decision Making and Nursing Research, Rikshospitalet-Radiumhospitalet Medical Center, Norway.*
[d]*School of Library and Information Studies, University of Wisconsin-Madison, U.S.A.*

Abstract

Through personal health record applications (PHR), consumers are gaining access to their electronic health records (EHR). A new challenge is to make the content of these records comprehensible to consumers. To address this challenge, we analyzed the text unit length, syntactic and semantic characteristics of three sets of health texts: clinical reports from EHR, known difficult materials and easy-to-read materials. Our findings suggest that EHR texts are more different from easy texts and more similar to difficult texts in terms of syntactic and semantic characteristics, and EHR texts are more similar to easy texts and different from difficult texts in regard to text unit length features. Since commonly used readability formulas focus more on text unit length characteristics, this study points to the need to tackle syntactic and semantic issues in the effort to measure and improve PHR readability.

Keywords:

consumer health, readability, personal health record, consumer health vocabulary, natural language processing

Introduction

Increasingly, consumers are taking an active role in their own health care by accessing and contributing to their personal health records (PHR). This role has become widely recognized by health care organizations and policy makers. One major source of PHR content is the institutional electronic health records (EHR), which are complex documents created by health care professionals for medical, legal, financial and administrative purposes. The organization, syntax, vocabulary and underlying conceptual knowledge employed by medical records may not be easily comprehended by lay people. For consumers with an average level of health literacy, understanding EHR content is challenging: Is "negative x-ray finding" good or bad? What does "FH of MI" mean? Which section(s) of the discharge summary describe my treatment plan?

For the PHR to fully realize its potential in helping consumers to manage complex health data and to facilitate informed decision making and self-care, its content should be easily understandable to consumers. A prominent panel of fellows of the American College of Medical Informatics recently published a white paper [1] stating "In order to be useful to the patient, the PHR must present data and accompanying tools in ways that enable the individual to understand and to act on the information contained in the record…..Both terminology and data presentation must be adapted to the individual using the PHR, so that they realize optimal benefits."

Considerable readability issues exist for today's PHRs which typically contain selected portions of an EHR and are aimed at a rather educated user group [2, 3]. The need to make the EHR information comprehensible will be more critical as an increasingly diverse patient population gains access to increasingly comprehensive records.

We have embarked on a project to translate EHR information into intelligible structure and plain language for PHR users. The goal of translation requires us to first understand and measure the readability of EHR information. This paper presents an analysis of EHR text characteristics as one of the initial steps toward text translation.

Background

Although more than a few health-specific literacy tests such as the Test of Functional Literacy in Adults (TOHFLA) have been developed [4], practically no health-specific readability measure is available. Recognizing the potential limitations of existing general-purpose readability measurements, we and other researchers began to examine various characteristics of health texts.

Our previous studies focused on the vocabulary aspect and resulted in the development of term and concept familiarity estimation methods [5]. In evaluation studies, our predicted term and concept familiarity was shown to be well correlated with actual consumer vocabulary knowledge and comprehension [6, 7] and outperformed the word length and word list techniques employed by the general-purpose readability formulas [5].

In a 2006 report, Rosemblat and colleagues examined what text features health communication experts use to determine the readability of consumer-oriented health texts [8]. The two significant factors they identified were "vocabulary" (i.e. number of words that are likely to be familiar to readers) and "main point" (i.e. ability of readers to identify

and understand the "take home" message). In this study, the presence and absence of these factors in the texts were established subjectively by the experts.

Also in 2006, Leroy and colleagues published a study that analyzed and compared the text characteristics of four types of documents: easy and difficult WebMD documents, patient blogs, and patient educational material, for surface and content-based metrics [9]. The easy and difficult WebMD documents were determined using the Flesch-Kincaid formula. They found a number of syntactic and semantic similarities and differences: for example, the easy WebMD pages are the most similar to patient blogs in terms of vocabulary difficulty.

No previous study has examined the readability-related characteristics of clinical reports in EHR systems, though a study by Chapman et al. did apply the Flesch-Kincaid formula to a set of dictated and transcribed x-ray reports [10]. While it was not the authors' finding, we observed from the results of Chapman's study that the readability measure (Flesch-Kincaid) was greatly underestimating the difficulty of these reports: the average grade level of these reports was reported to be 7.6. Based on our experience of natural language processing (NLP) of radiology reports, they are often difficult for non-clinician researchers with graduate school education (equivalent to grade level 18 and above) to comprehend.

Materials and methods

Materials

We collected three sets of health documents: EHR reports, difficult texts and easy texts.

The first set contains 40 EHR reports randomly selected from the clinical data repository of the Brigham and Women's Hospital and Massachusetts General Hospital (Boston, MA, U.S.A.). We retrieved 10 outpatient clinic notes and 10 discharge summaries from each institution. The reports cover topics such as chief complaint, history of illness, laboratory finding, treatment, and discharge plan. The medical diagnoses which appeared in the reports included common disease such as asthma, diabetes mellitus, pneumonia, and osteoarthritis. The average length of the documents is 3374 characters.

The second set is 40 abstracts of scientific journal papers randomly retrieved from MEDLINE (www.pubmed.org). The majority of journals indexed for MEDLINE are intended for a readership of researchers and clinicians, and typically require substantial background knowledge in specialty areas (e.g. molecular biology or nephrology) to understand. The abstracts, thus, are good examples of materials that are difficult for lay health consumers. This set of documents included various topics such as abdominal pain, asthma, hypertension, and paranoid schizophrenia. The average length of the documents is 1801 characters – abstracts are short by nature.

The third set is a convenience sample of 40 easy-to-read documents. We collected 27 (self-labeled) easy-to-read documents from multiple high-qualify consumer health Web sites: 21 from the Food and Drug Administration (www.fda.gov), 4 from the National Institute of Mental Health (www.nimh.nih.gov), and 2 from the National Institute on Alcohol Abuse and Alcoholism (www.niaaa.nih.gov). We also selected 13 records from the Reuters Health (http://www.reutershealth.com). The topics covered by these easy-to-read materials varied as well, including allergy, heart attack, breast feeding, alcoholism, and depression. The average length of the documents is 4101 characters.

Methods

Each document was processed by HITEx – a suite of open-source NLP tools that we have developed [11]. Each document was tokenized, split into sentences, and had part-of-speech (POS) tags assigned. Noun phrases were subsequently extracted and mapped to the Open-Access Collaborative (OAC) consumer health vocabulary[1].

For each parsed document, we first calculated the total number of characters, words, sentences and paragraphs. We considered a word to be any token that does not contain punctuation symbols. Paragraphs were defined depending on the document style: In the EHR reports we used, paragraphs are separated by a blank line; in the easy text sample, they are marked by line breaks. We then calculated the average word length (i.e., number of characters per word), average sentence length (i.e., number of words per sentence), and average paragraph length (i.e., number of sentences per paragraph).

Next, we calculated the frequency distribution of POS categories in each document. For the purpose of statistical analysis, some less frequent POS categories were merged (e.g., all punctuation categories were merged into one), reducing the total number of categories from 30 to 13.

Thirdly, we calculated the average term and concept familiarity scores for each document. These scores were obtained from the OAC consumer health vocabulary[1]. The OAC vocabulary provides three scores: a frequency-based term score (derived from term occurrence data), a context-based term score (derived from term co-occurrence data) and a context-based concept score (derived from concept co-occurrence data). The term scores reflect the string (surface)-level difficulty for consumers and the concept scores reflect the concept-level difficulty for consumers [6]. The scores have the range between 0 and 1, with 1 indicting perfect consumer familiarity (i.e., the easiest) and 0 indicting complete consumer unfamiliarity (i.e., the most difficult). We used the scores to gauge the semantic complexity of the contents.

Some terms did not map to OAC and not all OAC terms had the three scores assigned yet. When we calculated the weighted averages of the scores, these out-of-dictionary terms and terms with missing scores were excluded.

Finally, we calculated the Flesch-Kincaid grade level [12] for every document using a Microsoft Word™ built-in function.

In statistical analysis, mean and 25% and 75% quintiles of each text characteristic were first calculated. We then tested whether the EHR set shares the same text character-

1 More detailed explanations and related publications of this vocabulary and the term/concept familiarity scores it provides can be found on the Consumer Health Vocabulary Initiative's Web site (www.consumerhealthvocab.org).

istics with either the difficult or the easy document set. The distributions of the characteristics were examined using the Shapiro-Wilkins W test ($p < .001$) to assess normality. When distributions are not normal, differences in distributions were tested using the Wilcox rank sum test. The text characteristics with normal distributions were tested for differences in means using the t-test.

Results

The text characteristics of the three sample sets (EHR, difficult, and easy text) were different in many aspects. The mean and 25%/75% quantiles of the text characteristics are reported in Table 1-3.

Table 1 - Means and differences in text unit length characteristics

	EHR	Easy	Difficult
Average # of Characters per Word^	4.97 (4.78, 5.23)	4.71* (4.40, 5.02)	5.53* (5.28, 5.83)
Average Num of Words per Sentence	13.46 (10.32, 16.35)	16.93 (12.38, 21.89)	17.60** (14.53, 20.41)
Average Num of Sentences per Para–graph	3.57 (1.96, 3.41)	1.84 (1.45, 2.0)	15.50** (13.0, 18.0)

Lower and upper quantiles are in parentheses.
* compared to EMR, $p < .05$ ** compared to EHR, $p < .0001$
^ tested for means using t-test.

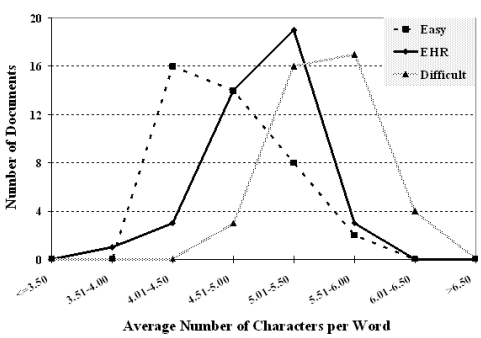


Figure 1 - Distributions of the num of characters per word

On the text unit length level, the EHR sample falls between the easy and difficult texts in terms of word length (Figure 1); it has the shortest sentence length (Figure 2), which is not statistically different from that of the easy texts; it also has very few sentences per paragraph (Figure 3), which is not statistically different from that of the easy texts while being very different from that of the difficult texts. The short average word and sentence lengths of the EHR sample are largely due to the use of abbreviations and incomplete sentences.

On the syntactic level, the EHR sample differs from the easy texts statistically in most of the POS categories (Table 2). It does have some similarity with the difficult texts, for example, both have high proper noun usage and less verb and adverb usage.

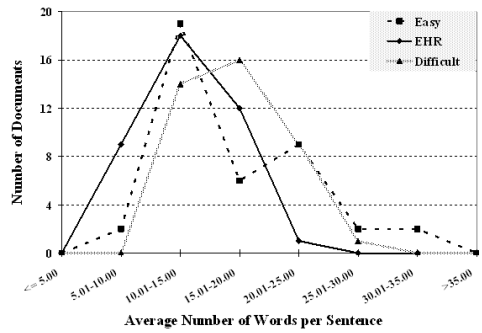


Figure 2 - Distributions of the number of words per sentence

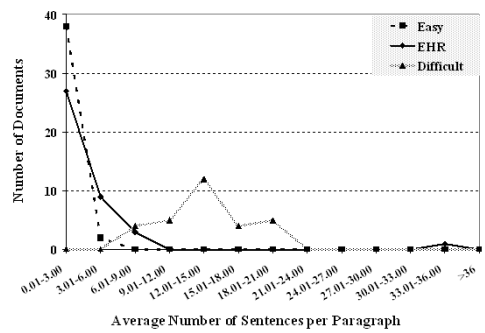


Figure 3 - Distributions of the number of sentences per paragraph.

Table 2 - Means and differences in syntactic characteristics (parts of speech categories per sentence)

	EHR	Easy	Difficult
Verb	1.62 (1.31, 2.01)	2.70** (2.09, 3.21)	1.93* (1.41, 2.29)
Noun^	3.30 (2.27, 4.08)	4.75** (3.29, 6.08)	5.48** (4.47, 6.19)
Proper Noun	3.32 (1.62, 4.50)	1.27** (0.51, 1.71)	3.01 (2.0, 4.0)
Punctuation	2.44 (1.77, 2.70)	2.37 (1.69, 3.17)	3.39** (2.63, 4.0)
Pronoun	0.36 (0.15, 1.56)	0.65** (0.41, 0.87)	0.09** (0, 0.15)

	EHR	Easy	Difficult
Adverb	0.33 (0.16, 0.51)	0.58** (0.44, 0.67)	0.40 (0.24, 0.47)
Adjective^	0.87 (0.54, 1.16)	1.41** (0.95, 1.67)	1.76** (1.43, 2.02)
Particle	2.02 (1.33, 2.53)	3.12 (2.16, 3.94)	3.09** (2.48, 4.70)
Determiner	0.03 (0, 0.06)	0.10** (0.04, 0.12)	0.05 (0, 0.07)
Proposition	0.02 (0, 0.04)	0.03 (0, 0.05)	0.02 (0, 0.04)
Symbol	0.09 (0, 0.08)	0.00** (0, 0)	0.06 (0, 0.06)
Modal	0.09 (0.02, 0.10)	0.32** (0.25, 0.39)	0.09 (0, 0.13)
Possessive	0.35 (0.06, 0.53)	0.47 (0.27, 0.62)	0.05** (0, 0.10)

Lower and upper quantiles are in parentheses.
* compared to EMR, p <.05 ** compared to EHR, p <.0001
^ tested for means using t-test.

Table 3 - Means and differences in semantic characteristics

	EHR	Easy	Difficult
Average Context-based Term Scores	0.60 (0.58, 0.66)	0.72** (0.68, 0.80)	0.62 (0.53, 0.70)
Average Frequency-based Term Scores^	0.63 (0.59, 0.67)	0.77** (0.74, 0.80)	0.67* (0.63, 0.71)
Average Context-based Concept Scores^	0.66 (0.65, 0.68)	0.73** (0.71, 0.75)	0.68 (0.65, 0.71)

Lower and upper quantiles are in parentheses.
* compared to EMR, p <.05 ** compared to EHR, p <.0001
^ tested for means using t-test.

On the semantic level, EHR's mean familiarity scores are the lowest in all three metrics (Table 3). The scores range from 0 to 1, terms/concepts with lower scores are considered to be less familiar to consumers and more difficult. Statistically significant differences were found between EHR and easy texts in every metric, while EHR and difficult texts only differ on the frequency-based term scores. The score distribution curves of the EHR and difficult texts overlap to a large extent (Figures 4-6).

It is a common belief that the EHR is not easy for consumers to understand and there is ample empirical evidence supporting the view [1]. The comparison of EHR text characteristics with the characteristics of the difficult and easy health text samples suggest that the syntactic and semantic characteristics are the key to explain EHR's low readability. On the other hand, the text unit length features (word, sentence and paragraph lengths) failed to account for the difficult of EHR texts: EHR texts are similar to easy materials regarding these features.

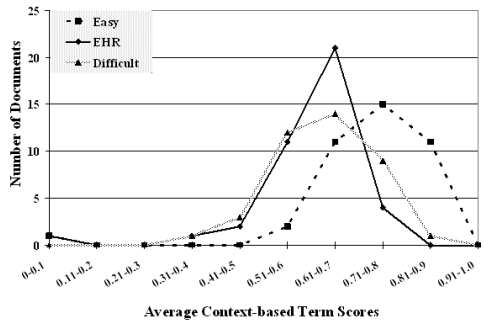


Figure 4 - Distributions of the average context-based term scores

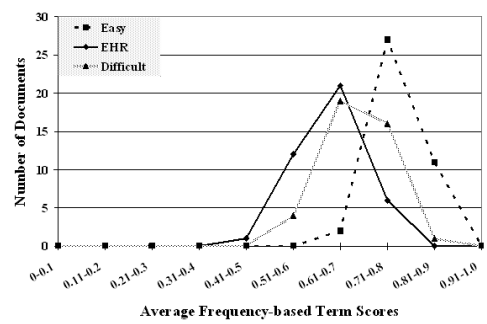


Figure 5 - Distributions of the average frequency-based term scores

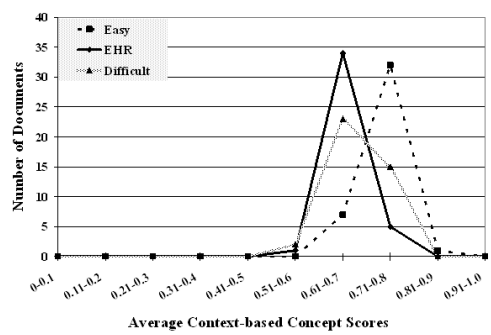


Figure 6 - Distributions of the average context-based concept scores

Table 4 - Means and differences in readability scores

	EHR	Easy	Difficult
Flesch-Kincaid Grade Levels	9.68 (8.55, 11.50)	8.23* (5.65, 12)	11.98** (12.0, 12.0)

The Flesch-Kincaid levels of the three samples also differ statistically (Table 4). The average grade of 8.23 is probably an accurate assessment of the easy texts. The 11.98 grade level assigned to the difficult texts is an underestimation, however, can be partially blamed on the abstract style. The 9.68 grade level of EHR, though, is clearly inaccurate.

Discussion

Although the clinical reports in EHRs are primarily written and read by health professionals, consumers are gaining access to them through the proliferation of PHRs and the readability of EHR reports for consumers has been recognized as a problem. This paper presents an analysis of the text unit length, syntactic, and semantic characteristic of EHR texts, and their implications for PHR readability.

Our statistical analysis indicates that EHR texts are more different from known easy texts and more similar to known difficult texts on the syntactic and semantic levels, while EHR texts are more similar to easy texts and different from difficult texts on the text unit length level. On the other hand, the commonly used readability formulas focus on text unit length rather than syntactic and semantic features, which we believe is the main cause of Flesch-Kincaid formula's inaccurate assessment of the difficulty of the EHR reports. To measure and improve EHR readability for the PHR audience, syntactic and semantic characteristics must be taken into consideration.

One may argue that it is a bit farfetched to compare the text unit length characteristics of EHR and easy texts, since they are obviously different types of documents. Please note that most of the readability formulas which are commonly used by biomedical researchers for a wide range of text materials are based on text unit length. The similarity between EHR and easy texts in terms of unit length points to the limitations of the unit length-based measurements.

Some limitations of this study are: Although the text samples we used are comparable in size to a couple of the related studies described in the Background section, they are small. MEDLINE abstracts are good examples of difficult health texts; however, they do not provide a diverse representation of difficult health texts. In follow-up studies, full length articles will be used as well.

Our results suggest that there is a need for an EHR specific readability measure and we plan to develop such a metric. We are also interested in exploring the role syntactic and semantic characteristics play in other health texts that consumers are exposed to and validate our findings through user studies. The ultimate goal of our research is to translate EHR texts into a lay-friendly language for PHRs.

Acknowledgments

This work is supported by the NIH grant R01 LM07222 and by the Intramural Research Program of the NIH, NLM.

References

[1] Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc 2006;13(2):121-6.

[2] Hassol A, Walker JM, Kidder D, Rokita K, Young D, Pierdon S, Deitz D, Kuck S, Ortiz E. Patient experiences and attitudes about access to a patient electronic health care record and linked web messaging. J Am Med Inform Assoc 2004;11(6):505-13.

[3] Winkelman WJ, Leonard KJ, Rossos PG. Patient-perceived usefulness of online electronic medical records: employing grounded theory in the development of information and communication technologies for use by patients living with chronic illness. J Am Med Inform Assoc 2005;12(3): 306-14.

[4] Baker D, Williams M, Parker R, Gazmarian J, Nurss J. Development of a Brief Test to Measure Functional Health Literacy. Patient Education and Counseling 1999;38:33-42.

[5] Zeng Q, Kim E, Crowell J, Tse T. A Text Corpora-Based Estimation of the Familiarity of Health Terminology. Lecture Notes in Computer Science 2005;3745 / 2005:184-192.

[6] Keselman A, Tse T, Crowell J, Browne A, Ngo L, Zeng Q. Assessing Consumer Health Vocabulary Familiarity: An Exploratory Study. In: MEDNET; 2006; Toronto, CA; 2006.

[7] Keselman A, Massengalea L, Ngo L, Browne A, Zeng Q. The Effect of User Factors on Consumer Familiarity with Health Terms: Using Gender as a Proxy for Background Knowledge about Gender-Specific Illnesses. In: International Symposium on Biological and Medical Data Analysis; 2006; Thessaloniki, Greece; 2006. In press.

[8] Rosemblat G, Logan R, Tse T, Graham L. How Do Text Features Affect Readability? Expert Evaluations on Consumer Health Web Site Text. In: MEDNET; 2006; Toronto, CA; 2006. In press.

[9] Leroy G, Eryilmaz E, Laroya BT. Health Information Text Characteristics. In: American Medical Informatics Association (AMIA) Annual Symposium; 2006; Washington DC; 2006. In press.

[10] Chapman WW, Aronsky D, Fiszman M, Haug PJ. Contribution of a speech recognition system to a computerized pneumonia guideline in the emergency department. Proc AMIA Symp 2000:131-5.

[11] Zeng QT, Goryachev S, Weiss S, Sordo M, Murphy SN, Lazarus R. Extracting principal diagnosis, co-morbidity and smoking status for asthma research: evaluation of a natural language processing system. BMC Med Inform Decis Mak 2006;6:30.

[12] Zakaluk BL, Samuels SJ. Readability: Its Past, Present, and Future: Intl Reading Assn; 1988.

Address for correspondence

Qing Zeng-Treitler, Ph.D.
Decision Systems Group
Brigham and Women's Hospital
Harvard Medical School
75 Francis Street Boston, MA 02115
Email: qzeng@dsg.harvard.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Generic Screen Representations for Future Proof Systems – Is It Possible? Two-model approach to a generic GUI

Helma van der Linden[a], Thilo Schuler[b], Rong Chen[c], Jan Talmon[a]

[a] *Medical Informatics, University Maastricht, The Netherlands*
[b] *Department of Medical Informatics, University of Freiburg, Germany*
[c] *Department of Biomedical Engineering, Linköping University, Sweden*

Abstract

Semantic interoperability should not only cover system interpretation of incoming information, but should be extended to include screen representation. This article describes a two-model approach to generate a screen representation for archetype-based information, which is inspired by the two-model approach used by openEHR for their archetypes. It provides a separation between software-related display knowledge and domain-related display knowledge and is designed with reuse of components in mind. This approach leads to a flexible GUI that can adapt not only to information structures that are not predefined within the receiving system and display them in a meaningful way, but also to novel ways of displaying the information.

We are working on a proof of concept implementation to validate the approach.

Keywords:

computerized medical record, electronic health record, user-computer interface, openEHR, archetypes, HL7

Introduction

Patient mobility is increasing over the last decade varying from "shopping around for care" in different health care organizations to prolonged stays abroad with the increasing need of care. This results in fragmented patient-related health information distributed across different systems.

A current approach for future health information systems is to create a virtual health record that integrates the fragmented information by connecting distributed systems and presenting them as one. The underlying principle is information exchange based on standardized messages. The two major approaches in this respect are HL7 v3[1] [1] and CEN/TC251 13606 [2, 3] combined with openEHR archetypes[2] [4].

With ongoing efforts towards harmonization of the best of both frameworks it will be possible to see virtual health records come into reality within some decades.

In this article we assume the existence of such an environment where virtual health records exist and information exchange is not limited to the systems of a single organization, but expanded to incorporate other organizations on a regional or maybe even global scale. In this environment it is possible to retrieve or receive patient information of which the structure has not been known before. Indeed, as advocated by openEHR, true future-proof electronic health record systems will be able to accommodate new medical concepts without the need for redevelopment. The key idea behind archetypes is to express new information structures as a combination of predefined classes.

The openEHR Foundation has currently the only architecture that allows handling of unknown information structures. We therefore focus on openEHR archetypes in this article. This does not imply that archetypes are the only means of exchanging information.

Scenario

A GP suspects that the patient suffers from a hereditary disease and refers the patient to the genetics clinic where tests will be done to confirm or reject his suspicion. Unfortunately, his suspicion is confirmed and the GP receives a discharge letter that contains a summary, the lab results and a family tree.

When we assume that the discharge letter is a structured message containing the data structures with the relevant data rather than a formatted display document, the question arises how the information of the discharge letter should be displayed on the GP's screen and in particular the information that is normally not part of the GP's system (i.e. the family tree).

The article discusses an approach to display new information structures on a user's screen using as much display knowledge as available.

Background

ISO 18308 defines semantic interoperability as the ability for information shared by systems to be understood at the

1 In this article HL7 will refer to the new v3 standard in its latest form.
2 In this article we will refer to the CEN 13606/openEHR archetypes as openEHR or archetypes for brevity.

level of formally defined domain concepts so that the information is computer processable by the receiving system [5].

Both HL7 and openEHR support semantic interoperability at two levels: at the data structure level and at the domain level.

At the data structure level, medical concepts are described using predefined data structures. This ensures that the information exchanged is complete (i.e. it contains all relevant data and metadata) and can be parsed, stored and subsequently retrieved. At the domain level metadata such as the code and coding scheme are used to avoid ambiguity in understanding.

The PropeR project has revealed that semantic interoperability is not simply a matter of interoperability between systems, but also between user and system. To refer to the scenario, even if the GP's system is capable of storing and subsequently retrieving the fully structured family tree, if there is no suitable screen representation, it is very difficult for the GP to correctly interpret the information.

Van der Meijden [6] and van Ginneken [7] have already discussed the difference between data entry and data retrieval with respect to the screen representation. Since it is logical to assume that data entry is only done in the local system, the issue of undefined data structures does not occur during data entry as we may assume that the local system is designed to support the specific user tasks in the application domain. The approach we present here will primarily be focused on data consultation and not on data entry.

Methods

In the context of the PropeR project [8, 9] we built a web based EHR system based on a simplified version of archetypes. We focused on the implementation of a domain-agnostic system and the strict separation between archetypes and screen representations.

We followed a similar approach by researching the feasibility of generating a GUI based on openEHR archetypes [10].

The lessons learned in both projects were combined to develop a more generic approach that can handle the situation we discussed before.

Results

Presentation level interoperability

In our view displaying information in an unambiguous way that supports the user's work processes, requires three types of knowledge:

- Knowledge of the information to display;
- Knowledge of the way a user is accustomed to view information;
- Knowledge of the device that is used to display the information.

Information-related presentation knowledge

At the lowest level this refers to the display of the data types that are used to construct the information structure: numbers are displayed differently than text. This is however, not sufficient. Even the example of a simple blood pressure shows that a higher level of knowledge is necessary to correctly display a blood pressure; that is in the common form of two numbers separated by a slash. A graphic tree form would best represent a family tree.

Localized presentation knowledge

Displaying information can be subject to local customs, varying from the local language and the local date format to preferred units (e.g. mg/dl vs. µmol/l) and coding schemes. There are also personal differences in what the best way of information presentation is with different reasons such as learned behavior or different cognition strengths (e.g. visual, textual).

Device-related presentation knowledge

The current trend towards ubiquitous computing has produced a large range of devices capable of sending and retrieving information ranging from desktop computer and laptops to tablet pcs, pda's and smartphones. While each modern model contains a webbrowser, and thus an abstraction from the underlying device, the supported functionality and the screen size places extra constrains on the presentation.

These different types of knowledge are often hard coded into the GUI of the client application. This makes it very hard to display incoming information from a different domain.

A two-model approach to generic GUI generation

Given the premise that future-proof systems are also capable of displaying information from other domains, it is necessary that these systems contain domain-agnostic screen representational functionality.

From the PropeRWeb application we learned that screen representation knowledge, however low-level, should not be incorporated in the archetype definition [11]. Not only does it introduce two different kinds of knowledge (medical domain knowledge and presentation knowledge) in a single model, but it is also common knowledge that a single data type, especially numerical, can be displayed in different ways, for example as a single number or as a table or graph.

In our approach we distinguish two models: a display oriented model (the GUI model) that defines widgets as screen presentation units and a domain oriented model (the content model) that defines content units which create meaningful presentations using widgets. The first model is the realm of the GUI designer, while domain experts use the second model.

Localized presentation knowledge is defined in profiles and views. This results in four sets of presentation units, which are shown in figure 1: widgets, content units, views and profiles.

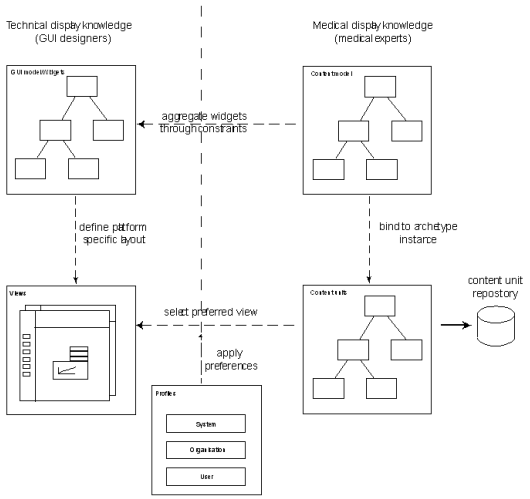


Figure 1 - Two-model approach to a generic GUI

All presentation units follow object-oriented design in which a specific unit inherits characteristics from a more generic unit. This improves consistency and flexibility. It is also valid to define multiple units for the corresponding information unit.

GUI model

The building blocks of a GUI are widgets. A widget is a display unit that contains presentation knowledge for a single data type. These widgets can be mapped to classes in the Reference Model of openEHR. Two types of widgets exist: data-oriented widgets such as "text", "image" and "number" and layout-oriented widgets such as "list" and "table".

The GUI model defines generic widgets that are converted to specific versions in the underlying system by using views.

Content model

Content units are defined using a content unit definition language. They are a semantic aggregation of widgets. They can be regarded as the display counterpart of archetypes. Content units, like archetypes can include other content units. Content units specify the binding to the information in the archetype instance as well as an established layout, such as the X/Y format of a blood pressure. Note that this layout only specifies relative positions of the included widgets and/or content units.

At the top level a content unit matches a COMPOSITION. These high-level content units are called *documents*. Different documents can be designed to reflect the differences in users' roles.

Content units can also include calculations, e.g. the total score of a test. They can also include normal ranges for semantic interpretation of the value. For example: the value of a body mass index can be color-coded based on the semantic interpretation (e.g. "normal" is green, "obese" is red).

Like archetypes content units are stored in a content unit repository. Since they are a semantic, platform independent representation of an archetype, they can be shared in the same way archetypes are sharable.

Views

Views can be regarded as implementations of content units customized for the device or application that will be displaying the information. Views can also include other GUI artifacts such as navigation bars.

A view is focused on presentation of the content and therefore part of the GUI designers' realm.

Profiles

The focus of profiles is the conversion of the information to match the user's expectations and thus avoid interpretation errors.

A profile contains preferences at various levels that modify the presentation of the information. There are three levels:

- *System level*. This level contains generic preferences that should always be applied, e.g. language, date format, metric vs. imperial system etc.
- *Local level*. This level contains generic preferences that are organization or location specific and are more domain-related. These preferences include preferred units and preferred terminologies.
- This level can also include role-based preferences that refer to role-based documents.
- *User level*. This level contains specific user related preferences that can modify the preferred view for a certain type of COMPOSITION e.g. if the user prefers graphs to tables.

Presentation generation

The process of generating the presentation is based on the pipeline concept. A pipeline can be compared to an assembly line where material arrives in a certain form, which is then processed by various stages along the line and finally delivered as a complete product. Adding or removing stages delivers a different product without affecting the other stages. A successful implementation of the pipeline concept can be found in Apache Cocoon [12].

A pipeline offers a component-based approach to the transformation of information. By adding or removing transformation components, the end result can change without affecting the other components. Different pipelines can implement different functionalities while sharing components.

In our approach an openEHR composition enters the pipeline. This composition can be the result of a query for information or the result of a notification of new information. In both cases the composition can contain instances of unknown archetypes.

Transformations handle device selection, presentation units selection, profile application and the final rendering of the selected view.

Figure 2 shows the generation process.

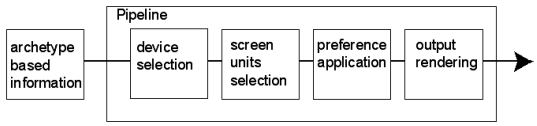


Figure 2 - Generation of view

There are several advantages to this solution:

- First and foremost this approach offers the flexibility of defining specific, optimized screen representations for known information structures, while providing the means to generate useable screen representations of unknown information structures.
- By separating the various types of presentation knowledge into distinct models, it is possible to separate pure GUI knowledge from medical display knowledge, thus honoring the two-model approach and promoting reuse.
- This approach is flexible enough to build a role-based GUI. Nurses and physicians can see the same information but optimally presented for their specific needs, while the only difference in development might be a document definition.
- The evolution of medical knowledge will always create new data types and new archetypes, which would lead to new display representations. Our approach ensures that as much of the available display knowledge can be reused.
- New and novel ways to view EHR information are a topic of ongoing research. By adhering to the proposed approach these views can benefit from the available display knowledge that is already expressed in content units. [13]
- Both HL7 and openEHR archetypes are using a limited set of predefined data types with an ongoing effort to harmonize the sets between the two parties. By describing one or more widgets for each data type it should be possible to provide a meaningful display of the information, without incorporating presentation knowledge in the information structure. This means the current archetype or message specifications need not be extended and the number of widgets is not very large.
- The information can be converted to match local and user preferences such as preferred coding scheme, language, units and more.
- A fallback mechanism is used to select a more generic representation in the absence of a specific one.
- Standardized content units could be shared between systems; the same way archetypes can be shared. This increments the intelligence and usability of the system.
- The pipeline approach not only allows reuse of components, but also offers flexibility in adding functionality by a simple addition of pipelines.

This approach complements the openEHR architecture where templates are used to create a higher-level composition by constraining and ordering archetypes. In contrast, the openEHR templates are used to *create* an information structure, while the approach proposed in this paper is used to *display* the information.

HL7 focuses on message exchange only and therefore considers this problem to be part of the receiving system's domain. However, given the similarities in structure between archetypes and messages we believe this approach is equally useful in that realm.

There are also disadvantages:

- Higher-level, specific views can only exist for predefined information. New information or information from different domains will fall back to a more basic representation.
- A repository, equal to that for archetypes, is necessary for the various presentation units.
- A mechanism for retrieving an appropriate screen representation for the current archetype is necessary, since the most appropriate selection is based on multiple parameters, described earlier, that cannot be stored in the archetype instance.

The advantages of having flexible GUI interfaces outweigh the disadvantages. By incrementally defining screen representations that can be built on top of each other, there is less duplication of work in building a GUI. A higher-level screen representation allows the user to better interpret the presented information thus leading to more efficient and more reliable information exchange. Screen representations for new information structures can be added to the system without major redevelopment of the application.

Currently we are working on a proof of concept using the Apache Cocoon web application framework [12] to build a web application that can display instances of various openEHR archetypes based on the approach described here.

The Apache Cocoon web application framework is a generic open source framework that is heavily based on the concept of separation of concerns to define strict distinctions between model and view. It implements the pipeline concept and also provides a set of generic widgets and an XML-based language to define what we call views. Since Cocoon excels in processing XML it is a good candidate to build a generic generated web based GUI using the approach that we have presented before. A first version will be presented in the openEHR workshop of Medinfo 2007.

Related work

A similar approach is developed by Ocean Informatics and implemented in their EhrView [14]. The EhrView application modifies the information through a series of XSLT stylesheets. These stylesheets are selected by matching the archetypes names in the composition. The matching process selects the most specific stylesheet available in a repository.

The EhrView application does not separate content related modeling from software related modeling and it is only defined for one type of device: a regular screen of a desk-

top or laptop pc. It also offers limited options to adjust to local or user preferences.

Fiala et al. [15] have described a component-based approach for adaptive web documents that influenced our approach. They too make a distinction between content-related and display related presentation knowledge and they also used the pipeline concept to define web document generation. However, their focus is on adapting information presentation to user preferences and devices. The information is known and defined in advance and there is no method to handle unknown information structures or describe conversions to preferred units.

Conclusion

Semantic interoperability does not stop when information from one system can be successfully understood and/or incorporated in another system. It is also necessary to provide a screen representation that gives the user of the receiving system a clear understanding of the new information.

In this article we described an approach that extends the two-model approach that is currently used by openEHR by a similar approach for the GUI.

We argued that this approach leads to a flexible GUI that can adapt to information structures that are not predefined and still display them in a meaningful, higher-level way. We are currently working on a proof of concept implementation.

Acknowledgments

We would like to thank the team from Ocean Informatics who kindly provided the information we needed. We would also like to thank Patrick Ahles for his valuable insights.

References

[1] Health Level 7 Inc. [last visited 21 Mar 2007]; Available from: http://www.hl7.org

[2] Health Informatics – Electronic health record communication - Part 1: Reference Model: CEN/TC251; 2005 March 2005. Report No.: prEN13606-1:2005:E.

[3] Health Informatics – Electronic health record communication - Part 2: Archetypes: CEN/TC251; 2005 December 2005. Report No.: prEN13606-2:2005:E.

[4] openEHR. [last visited 21 Mar 2007]; Available from: http://www.openehr.org

[5] ANSI. ISO/TS 18308 Health Informatics - Requirements for an Electronic Health Record Architecture: ISO; 2003 2003-04-23.

[6] van der Meijden MJ, Tange HJ, Boiten J, Troost J, Hasman A. An experimental electronic patient record for stroke patients. Part 2: system description. Int J Med Inform. 2000 Sep;58-59:127-40.

[7] van Ginneken AM. Considerations for the representation of meta-data for the support of structured data entry. Methods Inf Med. 2003;42(3):226-35.

[8] van der Linden H, Boers G, Tange H, Talmon J, Hasman A. PropeR: a multi disciplinary EPR system. Int J Med Inform. 2003 Jul;70(2-3):149-60.

[9] van der Linden H, Talmon J, Tange H, Grimson J, Hasman A. PropeR revisited. Int J Med Inform. 2005 Mar;74(2-4):235-44.

[10] Schuler T, Garde S, Heard S, Beale T. Towards Automatically Generating Graphical User Interfaces from openEHR Archetypes. Stud Health Technol Inform. 2006;124:221-6.

[11] van der Linden H, Grimson J, Tange H, Talmon J, Hasman A. Archetypes: the PropeR way. Medinfo. 2004;11(Pt 2):1110-4.

[12] Apache Cocoon Web Application Framework. [last visited 21 Mar 2007]; Available from: http://cocoon.apache.org

[13] Sundvall E, Nyström M, Forss M, Chen R, Petersson H, Åhlfeldt H. Graphical Overview and Navigation of Electronic Health Records in a prototyping environment using Google Earth and openEHR Archetypes. MedInfo 2007; 2007; Brisbane: accepted for publication; 2007.

[14] Ocean Informatics Pty L. EhrView. [last visited 21 Mar 2007]; Available from: http://www.oceaninformatics.biz/products/EhrView.html

[15] Fiala Z, Hinz M, Meißner K, Wehner F. A Component-based Approach for Adaptive, Dynamic Web Documents. Journal of Web Engineering. 2003 September 2003;2 (1&2):058-73.

Address for correspondence

Helma van der Linden
Medical Informatics
University Maastricht
POBOX 616
6200 MD Maastricht
The Netherlands
hvanderlinden@mi.unimaas.nl

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Speech Recognition in Dental Software Systems: Features and Functionality

Jeannie Yuhaniak Irwin, Shawn Fernando, Titus Schleyer, Heiko Spallek

Center for Dental Informatics, Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA United States

Abstract

Speech recognition allows clinicians a hands-free option for interacting with computers, which is important for dentists who have difficulty using a keyboard and a mouse when working with patients. While roughly 13% of all general dentists with computers at chairside use speech recognition for data entry, 16% have tried and discontinued using this technology. In this study, researches explored the speech recognition features and functionality of four dental software applications. For each system, the documentation as well as the working program was evaluated to determine speech recognition capabilities. A comparison checklist was created to highlight each program's speech functionality. Next, after the development of charting scripts, feasibility user tests were conducted to determine if performance comparisons could be made across systems. While four systems were evaluated in the feature comparison, only two of the systems were reviewed during the feasibility user tests. Results show that current speech functionality, instead of being intuitive, is directly comparable to using a mouse. Further, systems require memorizing an enormous amount of specific terminology opposed to using natural language. User testing is a feasible way to measure the performance of speech recognition across systems and will be conducted in the near future. Overall, limited speech functionality reduces the ability of clinicians to interact directly with the computer during clinical care. This can hinder the benefits of electronic patient records and clinical decision support systems.

Keywords:

Dental Informatics, Speech Recognition Software, User-Computer Interface, Medical Informatics Applications, Practice Management, Dental

Introduction

During care, dental clinicians have difficulty using a keyboard and a mouse, primarily because of infection control concerns but also because they are constantly using their hands for procedures and their office space and setup make it difficult to have the keyboard and/or mouse in close proximity. A solution to this problem, which is being employed in medicine [1-3], is the use of speech recognition applications to interact with the clinical computer. A recent study published on computing in clinical dental care offers insight into the adoption and use of speech applications in general dentistry [4]. Thirteen percent of all offices surveyed used speech input; however, 16% tried and discontinued using the technology [4]. Those who discontinued using speech did so because of technical problems with speech recognition (57%), lower efficiency compared to other data entry methods (13%), usability problems (9%), and other issues (22%) [4]. It is clear that there may be significant barriers to using the speech modules of current dental systems. Currently, dental speech applications typically implement command-and-control functionality as well as the transcription of free text [5]. The command-and-control functionality supports two types of activities: (1) navigating within the application (for instance to select a specific patient) and (2) entering structured data in text fields, list boxes, radio buttons and checkboxes [5]. Transcription is used primarily for dictating progress notes, surgical reports and similar textual information [5].

To date, there is no comprehensive overview or evaluation of the currently available speech recognition products in general dentistry. This study is significant because it is not only a comparative analysis of the four software applications, but it will also be an informative starting point for researchers interested in the benefits of speech modules as well as the development and design of future dental speech interfaces.

Materials and methods

Speech recognition features comparison

To compare the speech functionality of the software applications, the research team first acquired full working versions of the four practice management systems (PMS). Based on findings from a recent study, these four systems make up approximately 80% of the current practice management market in the United States [4]. The researchers reviewed each program's user manual to determine the system's speech features and functions. Then each system was installed according to its default installation configuration to allow the researchers to explore the programs' speech functionality. Researchers manually tested and used all components of speech functionality within each system. Lastly, the software vendors were contacted to answer any specific questions regarding the system's speech functionality and features. For example, a call to one of the companies was made when it was not clear what

speech engine was used in the system. A comparison checklist was created to highlight each program's technical details pertaining to speech input, including which features were present/absent as well as the program's limitations.

Feasibility user tests

In the second portion of this study, feasibility user tests were conducted to determine if the performance of speech recognition could be evaluated across systems. User tests have the potential to evaluate the efficiency, effectiveness, and user satisfaction of tested systems [6-8]. While four systems were evaluated in the feature comparison, only two of the systems were reviewed during the feasibility user tests. In our study, each user completed a task (efficiency and effectiveness) using one of the programs and then filled out a questionnaire (satisfaction) about their experience with the system.

To develop the task, a simulated intraoral patient record was created which contained a wide range of findings specific to testing the clinical speech capabilities of each system. Explicitly, the task was to chart 18 different hard and soft tissue findings via speech. During the feature comparison phase of this study, it became clear that to use the speech features of these systems, the user must have not only an in-depth working knowledge of the program, but must learn many specific and sometimes complex commands to interact via voice. Because of the complexity and specific command knowledge needed to complete the charting task via speech, participants could not be expected to complete the tasks on their own. Therefore, a step-by-step script was created for each of the two programs, which the participant used to test the system. As classic user testing does not involve using verbatim scripts [6], this unseen obstacle was one of the reasons feasibility tests were conducted on only two systems before performing an entire set of user tests. After development of the scripts, each was sent to their corresponding manufacturer to be evaluated for correctness and efficiency. Each manufacturer reviewed and edited its individualized script and then returned it to us with any changes in how the task should be completed via voice in their system.

To conduct the feasibility user tests, a computer (Windows XP, 1.5GHz Intel Pentium 4 processor, and 256MB of RAM) was equipped with an extra 80GB hard drive. This was necessary to store images of the machine with each software package installed. Norton Ghost 2003 (Symantec, Cupertino, CA) was used to make an image of the machine at baseline (with a fresh XP install). The first program and corresponding speech module were installed in default mode and configured with a patient family and a provider. Each system was used in its default installation configuration, that is, the speech interface for entering intraoral findings were not customized in any way in order to avoid "tuning" the program in preparation for entering the simulated patient. An image of the machine was then made with Ghost and the "first Program" image was stored. This was repeated for both systems. The result was three images, one of the machine at baseline, one for program one and one for program two. Before each user test the fresh installation image for the program being tested was restored. Using fresh images each time eliminated factors such as other user's voice profile existing, or other user's task results interfering with the next test.

Participants included three undergraduate students and one faculty member who work in our Center for Dental Informatics, at the School of Dental Medicine, University of Pittsburgh. The only criterion to participate in the study was a lack of experience with any of the speech features of the clinical charting interface of the PMS.

Four feasibility user tests were conducted, each user tested one of the two systems. Two users tested System One and two users tested System Two.

All programs that were tested required users to learn the speech aspect of the system via a brief training session. Each user was to test one of the two systems; therefore they only had to complete the one training session for their assigned program. The training sessions for the systems had minor differences, but generally, they were each approximately 20 minutes in length and required the user to read pre-determined words and sentences that appeared on the screen. Each user was supervised during his/her training session to assist with problems and questions. Assisting the participant during training ensured that the head-set microphone was adjusted properly and that the user was speaking optimally for the task. Successful completion of the training session was required to take part in the task evaluation.

To start the session, a background questionnaire was administered to each participant. The questionnaire was a modified version of a validated tool that measures dental students' use of, knowledge about, and attitudes towards computers [9]. In future user testing studies, the questionnaire will be used to determine any variance of the results based on users' age, sex, native language (English or non-English), prior computer experience, and affinity towards computers.

Next, each participant was randomly assigned to test one of the programs and asked to complete the required training for that program as described above. The participant was then asked to read the script to chart the 18 different hard and soft tissue findings via speech. Each participant was given the individualized script for the software program they were testing and asked to read the script verbatim. During the task completion, if the system's response resulted in being off script (e.g. the chart exited), the observer interrupted the participant, corrected the problem, and had the participant begin again either where they left off (if possible with out redoing steps) or start on the next finding. If the system did not respond at all, the participant was asked to repeat the command two more times (a total of three) and then asked to either move on to the next command, or if that was not possible, move on to the next finding (e.g. if the system does not select a tooth, the user will move to the next finding). If the system charted an error that did not result in being significantly off script (e.g. selecting the wrong tooth), the participant was asked to ignore the error and continue. When the task was complete, the user was asked to turn off the microphone.

During each session, two observers took hand-written notes and, to supplement data collection, the entire session for each participant was video recorded to capture the screen, including mouse clicks and audio.

Following completion of the script, each participant was asked to complete a user satisfaction questionnaire. The questionnaire contained 27 items answered via a 7-point Likert scale, and was based on the Subjective Assessment of Speech System Interfaces (SASSI) project [10]. The validated questionnaire by Hone and Graham is broken down into six main factors which can help predict a user's satisfaction with speech-based systems: system response accuracy, likeability, cognitive demand, annoyance, habitability, and speed [10].

To determine the efficiency and effectiveness of the systems, the following items were calculated: Time to complete the training and the time to complete the script (adjusted for off script actions). To evaluate effectiveness, all errors were recorded, including if there were any findings that were not charted because of detrimental errors (e.g. the chart closes). Errors included a detrimental error, which was defined as a time when the system's response resulted in the participant being completely off script. A repeated command error was recorded if the system did not respond and the participant had to repeat the command. In that case, each repeat was documented. A record was kept of which word was repeated. A wrong response error was documented if the system's response differed from the participant's input. Also, in this case, a record was kept out the system's response (e.g., did it choose another tooth, did it select distal instead of mesial, etc.). Lastly, if the system responded but the participant did not say anything, this was documented as an insertion error. Again, in this case, a record was kept about the system's response. To determine overall user satisfaction, the mean score for the user satisfaction questionnaires was calculated.

Results

Features and functions comparison

Table 1 shows which features and functions in each of the four systems could be completed via voice. Systems One and Two used Microsoft Speech Recognition Engine (Microsoft, Redmond, Wash.), whereas Systems Three and Four used the default speech engine installed on the computer as long as it had SAPI 4.0 or 5.0 program files. SAPI stands for Speech Application Programming Interface (Microsoft, Redmond, Wash). System One is the only program that allowed free text dictation into a "clinical notes" area, and this was done via a Dragon NaturallySpeaking Engine (Nuance Communications, Burlington, MA). The training sessions for all of the systems were documented to take approximately 5-10 minutes, and to use the free text dictation of System One, an extra 30 minutes of training was necessary. All of the systems had training sessions similar to the Microsoft Speech Recognition training. All of the systems allowed a user to complete extra training if necessary, and Systems Two and Three allowed the user to train with specific dental terms. Next, none of the programs allowed for naturally spoken text, i.e. all required specific speech commands; and the number of possible speech commands for each program were approximately 573 for System One, 140 for System Two, 41 for System Three, and 53 for System Four. Only System One allowed the use of the international communications alphabet (alpha, bravo) to assist with speech related interactions. All systems had the ability to provide audio confirmation of a given command, but only Systems Two and Three gave complete visual confirmation of commands. Systems One and Four did not provide visual conformation for some actions.

Table 1 - Functions that can be completed via voice

	Systems			
	One	**Two**	**Three**	**Four**
hard tissue charting	Yes	Some	No	No
periodontal charting	Yes	Yes	Yes	Some
dictate raw clinical notes	Yes	No	No	No
chart existing and proposed findings	Yes	Yes	No	Some
select tooth surface	Yes	Some	Yes	Some
select patient	Yes	Some	No	No
open chart	Yes	Yes	Some	No
select items from list via name shown	Yes	Some	No	No
navigate through chart ("next", "move down two", etc.)	Yes	Yes	Some	Some
use all displayed options and buttons	Yes	Some	Some	Some
access menus, buttons, pop-ups, and checkboxes	Yes	Yes	Some	Some
undo last command	Yes	Yes	Yes	Some
clear/delete entries	Yes	Yes	Some	Some
start and stop listening	Yes	Yes	Yes	Yes

Feasibility user tests

Initializing the computer with the images of each system and configuring each system with default settings and a generic patient family and provider took multiple attempts and a lot of time to perfect. Also, after the first user-test, it was determined that the format of the script - how the information was presented to the participant, as well as the general instructions - affected how well the participant could successfully complete the task. It was decided that because these were feasibility user tests, small aspects of the test would be changed after each test to eventually discover the optimum way to conduct the user tests in the future.

While user tests were done on only two systems, the scripts were for all four. These scripts which documented exactly how many and what steps were necessary to complete the task in each program, were in themselves major

findings for our study; see Table 2 for a comparison of the scripts. The total number of commands in each script were as follows: System One, 114; System Two, 92; System Three, 42; and System Four, 47; Systems Three and Four were only able to chart the periodontal findings via voice. Therefore, Systems One and Two had more than twice as many commands because they charted the hard tissue findings as well as the periodontal findings. In three of the scripts (Two, Three and Four) it was necessary for the user to utilize the mouse or keyboard at some point during the task completion, i.e. the charting could not be done via voice alone. It is also important to note that even though Systems Three and Four had no voice functionality for hard tissue charting, some hard tissue charting (charting missing teeth) was necessary to complete the periodontal part of the task.

Table 2 - Comparison of commands necessary to complete the charting task as documented in the scripts

	Systems			
	One	**Two**	**Three**	**Four**
total number of commands in script	114	92	42	47
total number of voice commands in script	69 (H) 45 (P)	41 (H) 45 (P)	0 (H) 38 (P)	0 (H) 39 (P)
total number of mouse/keyboard commands in script	0 (H) 0 (P)	5 (H) 1 (P)	4 (H) 0 (P)	3 (H) 5 (P)
percent completed with voice alone	100	93	90	82

(H) – Hard Tissue Charting, (P) – Periodontal Charting

The scripts demonstrate that in all systems, the commands are very specific and are directly comparable to using a mouse. For example, as opposed to being able to say "existing ML composite on tooth nine", the following had to be said (the quotation marks indicate that the command is spoken to the system): "select 9", "restorative", "move down 1", "ok", "mesial", "Lingual", "ok", "existing"(excerpt from script for System One).

Four user tests were conducted, two each on Systems One and Two. The first user tested System One, and because it was the first test, there were problems with the script format and the instructions, and there were many errors during the test. Therefore, changes for the next test were made and the data from the first user were discarded. The final results are based on three user tests, one with System One and two with System Two.

For System One, it took the user 11 minutes and 8 seconds to complete the training and 5 minutes and 20 seconds (adjusted) to complete the charting task (script). There were eleven repeated commands, the most frequent being the word "ok" which was repeated seven times. There was one detrimental error in which the system exited when the command was "existing". There were three wrong response errors, an example of one being when the system selected tooth 30 when the command was "select 3". Lastly there were three insertion errors during the pocket depth charting; numbers were inserted that were not said in a command.

For System Two, which two users tested, the average time to complete training was 9 minutes and 1 second and the average time (adjusted) to complete the script was 9 minutes and 13 seconds. There were an average of sixteen repeated commands, the most frequent being the word "ok", which needed to be repeated on average five times. There was one detrimental error in each test. For example, the user said "quick pick menu 11", and the system opened the patient history form. There were two wrong responses in each test, for example when the user said "3" for a pocket depth, the system thought it was selecting tooth three. There were no insertion errors in either test. The video recordings were used to verify the observed results.

As the user satisfaction questionnaire was based on a 7 point Likert scale, with the most positive answers being scored as a seven and the most negative answers being scored as a one the best possible score a system could receive is 189. The satisfaction score for System One (one user) was 104 and the average score for System Two was 77 (individual scores of 27 and 127).

Discussion

The results from this study show that current clinical software systems for dentists are attempting to accommodate speech recognition as a means of interaction, but the current systems have many limitations which may hinder their use. As shown in the feature comparison, speech functionality varies across all systems, with two of the systems not having the ability to complete hard tissue charting via voice, which was more than half of the common charting task. The fact that scripts had to be developed to conduct the user tests shows that these systems are not designed to be used without prior understanding of the software and the memorizing of or easy access to an enormous amount of specific terminology. The requirement of prior knowledge of the system is understandable due to the complexity of these programs; however if the commands and interaction with the system could be done with a more natural vocabulary, ease of use would be significantly improved. The scripts also show that three of the systems required at least one keyboard or mouse command to complete the charting task, which defeats the purpose of using speech recognition as a hands-free way to interact with the computer. If a clinician were to use one of these systems, she would still have to de-glove to interact with the computer, and the keyboard and mouse would still need to be easily accessible. This limited speech functionality has the potential to reduce the ability of clinicians to interact directly with the computer during clinical care, and may be the main reason 16% of dentists that tried using the technology later discontinued [4].

The feasibility user tests were only conducted with three participants; therefore it is impossible to make any generalizations or comparisons with the results. However, it is

clear that data can be collected in this manner for performance comparisons across the software packages. The average number of detrimental errors per system and the most frequently repeated words per system may provide insight into issues that can be resolved in future designs of dental speech systems. More training in specific dental vocabulary or more intuitive speech engines may be design suggestions resulting from future user tests.

Full scale usability testing will be conducted in the near future; however based on the feasibility tests, some minor changes in methodology will be made. For example, the individual charting tasks on each script will be grouped into findings for easier reading and a more straightforward task flow. Better instructions will be given to the users. Originally, because the participant was to read a script verbatim, very little instruction was given. After just one test, it was apparent more instructions and explanations were necessary. Next, the user satisfaction questionnaire will be modified. After reviewing the range of results from the questionnaire, it was discovered that having a script read verbatim is not conducive to enabling the participants to answer the questions asked on the validated satisfaction questionnaire. For the future tests, a new open-ended satisfaction survey will be develop to better asses user satisfaction based on the given task. Lastly, conducting and optimizing the user tests exposed the researchers to more speech related features and functionality of the systems that may have been overlooked during the first part of this study. Hence, more detail will be added to the feature and functionality comparison list after the completion of multiple user tests.

There are certainly limitations to this study. By using a verbatim script task, the dynamics of classic user testing are changed. More attention must be given to what is actually being compared across systems during this altered user test. Even though user tests will give a good indication of efficiency across all systems (i.e., which system completes the task in the least amount of time) unless these times are compared to charting the same information via a keyboard and mouse and ultimately to hand-written charting, we can never make accurate conclusions about efficacy compared to other data entry/retrieval methods.

Conclusions

This study shows that clinical speech functionality in current dental systems is somewhat cumbersome and poorly designed. This limited speech functionality has the potential to reduce the ability of clinicians to interact directly with the computer during clinical care. In the future, dentistry will see the influx and be able to reap the benefits of decision support tools and shared electronic medical records [11]. However, unless better speech functionality is implemented, the benefits and effectiveness of any electronic patient records and clinical decision support systems may be greatly impeded.

Acknowledgments

We gratefully acknowledge Thankam Thyvalikakath and Miguel Humberto Torres-Urquidy for their input on this project. We would also like to thank the NLM and NIDCR for their funding support 5T15LM007059-179.

References

[1] White K. Speech Recognition Implementation in Radiology. Pediatric Radiology 2005: 35(3): 841-846.

[2] Zafar A, Overhage M, McDonald CJ. Continuous speech recognition for clinicians. *J Am Med Inform Assoc* 1999: 6(3):195 -204.

[3] Grasso M. The long-term adoption of speech recognition in medical applications. In: Proceedings of the 16th IEEE Symposium on Computer-Based Medical Systems; 2003: p. 257-62.

[4] Schleyer T, Thyvalikakath T, Spallek H, Torres-Urquidy M, Hernanadez P, Yuhaniak J. Clinical computing in general dentistry. *JAMIA* 2006: 13(3): 344-352.

[5] Drevenstedt G, McDonald J, Drevenstedt L. The role of speech-activated technology in today's dental practice. *JADA* 2005: 136(2): 157-161.

[6] Nielsen J. Usability Engineering. San Francisco, CA: Morgan Kaufmann; 1994.

[7] Tomko, S. Speech Graffiti: Assessing the User Experience. CMU LTI Tech Report CMU-LTI-04-185 (2004). www.cs.cmu.edu/~stef/papers/mthesis.ps

[8] Harris TK and Rosenfeld RA. A universal speech interface for appliances. In: Proceedings of the 8th International Conference on Spoken Language Processing (ICSLP'04). Jeju Island, South Korea; 2004.

[9] Schleyer TK, Torres-Urquidy H, Straja S. Validation of an instrument to measure dental students' use of, knowledge about, and attitudes towards computers. *J Dent Educ*. 2001: 65(9): 883-91

[10] Hone KS, and Graham R. Towards a tool for the subjective assessment of speech system interfaces (SASSI). Natural Language Engineering 2000: 6(3-4): 287-303.

[11] Schleyer TK. Why integration is key for dental office technology. *J Am Dent Assoc*. 2004: 135(Suppl): 4S–9S.

Address for correspondence

Jeannie Yuhaniak, Department of Biomedical Informatics, Center for Dental Informatics, University of Pittsburgh, 333 Salk Hall, 3501 Terrace Street, Pittsburgh, PA 15261, USA. jeannie@dental.pitt.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Assessing the Impact of Recording Quality Target Data on the GP Consultation Using Multi-Channel Video

Maigaelle V. Moulene[a], Simon de Lusignan[a], George Freeman[a], Jeremy van Vlymen[a], Ian Sheeler[a] Andrew Singleton[a], Pushpa Kumarapeli[a]

[a] *Primary Care Informatics, Division of Community Health Sciences, St. George' - University of London, United Kingdom*

Abstract

Background: In the UK routinely collected computerized clinical data is used to assess progress towards financially incentivised quality targets for chronic disease management including hypertension. Objective: To develop a method for assessing the impact of recording quality target data in the clinical consultation. Methods: Raters were trained how to rate a multi-channel video of a simulated clinical consultation for interaction between actors, computer use, non-verbal communication. Results: 25% of consultation time is computer use and a median of 4 to 5 items were coded per consultation mainly items related to the hypertension quality target. Intraclass correlation coefficient showed good inter-rater reliability (>0.9; p<0.001). Conclusion: We have successfully piloted a novel technique for observing the influence of the computer on the consultation. Despite increasing computer use to record quality target data the over whelming proportion of the consultation remains doctor patient interaction.

Keywords:

video recording, consultations, computer, primary care, observation, medical records system, computerized.

Introduction

Chronic disease management is an increasing priority for general practice and in the UK financially incentivised quality targets have been introduced to raise the standard of chronic disease management [1]. As the population ages, people live longer but with an increasing burden of chronic disease such as hypertension. In the UK the Quality and Outcomes Framework (QOF) provides standards for the management of chronic disease. For example, in hypertension, blood pressure must be measured every 9 months and, ideally, kept below a target of 150/90. Performance of practices against these targets is measured using routinely collected computer data recorded as part of normal consultations. Reports are automatically generated from within the GP computer system based on the number of people with a diagnostic code for any of the chronic diseases and whether their disease management achieves the necessary target. Financially rewarded quality points are awarded based on achieving the target moderated by the prevalence of the condition in the practice compared with the national average. Management of hypertension attracts 105 points, each point worth £120 [177 Euros], to a practice of 10,000 patients, a total in excess of 18,000 Euros for each year the quality target is achieved. There is, therefore, considerable financial incentive for chronic disease data to be "coded," i.e. recorded as structured data, within the clinical consultation.

Video observation, using a single camera, is a well established method of assessing clinical competence in GP. Yet, there are still many limitations to single channel video, it is hard to interpret body language of the consulter or patient and it lacks information on how the computer is being used [2]. We have developed multi-channel methods to overcome the limitations of single channel video [3]. Three channel video uses several cameras capturing more detailed views of the doctor and the patient and includes a recording of how the computer is being used in the consultation, to allow computer analysis of the consultation and more in-depth analysis of the patient - doctor interaction. However, the shortcomings of this method were: the time taken and subjective nature of evaluation of the consultation; the difficulty in interpreting the patient's body language; and the high cost of making professional standard videos because analogue video at that time did not allow precise time sequence mapping. We have overcome the problems of cost using modern budget digital cameras; the addition of a fourth video channel looking at the patients body language; and, although we have made progress in using pattern recognition software to make automated records of the consultation this requires much more development if it is to record the subtleties of the clinical consultation [4]. We therefore used the four-channel refinement of our multi-channel video method [5] to analyse the time spent using the computer and how this might impact on the consultation; using simulated consultations for hypertension.

Methods

We carried out a literature review using standard bibliographic data bases restricting our search to articles published after 1990. We visited four GP practices to identify how clinical data are currently entered into the four main computer systems: EMIS PCS, EMIS LV, IPS Vision and iSoft Synergy. These four systems account for over

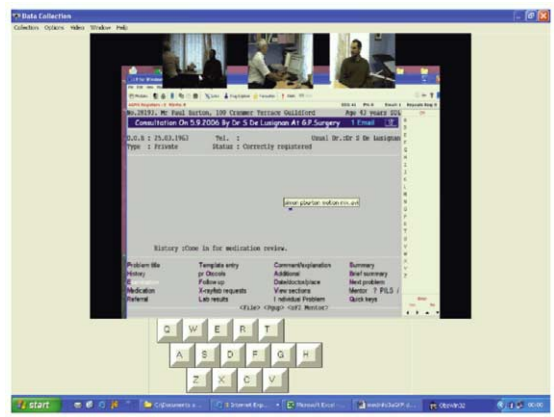
90% of GP clinical computer systems used in England. All these brands of GP computer system use the 5-byte version of the Read codes to record structured data. They also all have an integrated report tool which automatically reports anonymised data about their progress towards achieving quality targets. We spoke to experienced practitioners and discovered how chronic disease management data were entered into each system. We asked in detail about how the financially incentivised quality target (QOF) data are entered into each system and collected screen shots of each step in the process. We used the latter to train our raters how to recognize coding and recording of QOF target data.

We filmed simulated consultations with three of the GP computer systems: EMIS LV, EMIS PCS and IPS Vision in a real or simulated setting depending on the availability of the clinical computer system at our institution. We filmed three experienced GPs, all with over 15 years experience and one trainee GP. The GPs consulted the computer system they used in their clinical practice, limiting our choice to three out of the four systems initially observed; two of the GPs consulted using EMIS LV. The patients were simulated by members of staff experienced in role playing patients for clinical exams. We filmed a series of consultations to allow the GP to become familiar with the setup; however, we only used the consultation about hypertension for this analysis. Each clinician was presented with the same simulated patient for a review of their hypertension; their past history was identically loaded in each computer system.

We used three standard video cameras (Sony DCR HC45E DV) to record the consultation. One video camera recorded the doctor's head and upper body (to capture the clinicians direction of gaze and body language), the second the patient's upper body (we find that capturing the patients hands is important in interpreting their body language) and the third is an overview of the whole consultation (but excluding the examination couch). The fourth video channel was recorded using Camtasia screen capture software which records the computer screen and data entered into the clinical computer system. We mixed the separate video channels, so they could be displayed simultaneously, in-house. The separate video recordings were transferred from the mini DV cameras into Final Cut Studio editing software running on a G5 Mac to do this. The first step after transfer is to synchronise the images and audio feeds for each consultation. Next we created a layered video composition in Final Cut's compositing application (Motion). This composition displayed all the GP's computer usage and the camera angles in one screen. Finally this composition was rendered into AVI (Audio Video Interleave) format and recorded on to DVD.

We used ObsWin to analyse the consultations [6]. It integrates video files and uses keys on a keyboard where each key is a variable and acts as a 'stopwatch'. The raters had to watch the consultation video three times. We set up ObsWin so that one row of keys was used on each successive observation of the consultation (Fig 1). The first viewing used the lower row of keys to explore the interaction between the actors (including the computer): computer-doctor interaction (**Z,**), doctor-computer interaction (**X,**), doctor-patient interaction (**C**), patient-doctor interaction (**V**). The second run through used the second row of keys and measured data entry times: referrals (**A**), clinical coding excluding QOF target codes (**S**), quality target (QOF) coding (**D**), free text (**F**), prescribing (**G**), new prescribing needed to achieve a QOF target (**H**). The final screening used the top row of keys to measure body language occurring in the consultation: eye contact (**Q**), doctor looking at computer while speaking to patient (**W**), examination (**E**), patient speaking to doctor while doctor looking at computer screen (**R**) and Silent Time(**T**). Eye contact (**Q**) and silent time (**T**) are important non-verbal cues the former establishes rapport, shows the patient the doctor is engaged and trying to understand patient perspective, the latter also provides an important non-verbal cue [7]. We measured the time when the doctor was using the computer or looking at the computer while speaking (**W**) and when the patient speaks (**R**) to detect interference in the consultation. We could record examination (**E**) of blood pressure, but the examination couch was not under observation by our cameras thus other examinations (**E**) were not recorded and our simulated patients would have declined such examination. We derived information about the impact of coding from a combination of these variables. **C**+**V** represented the total verbal communication; **S**+**D** was the time spent on coding entries.

Figure 1 – ObsWin with integrated multi-channel videos



The raters were trained using an instruction manual [8] and special training videos. The raters were six volunteer biomedical informatics students from the authors' institution. They, MVM and an additional postgraduate student rated the consultations. Four or five raters rated every consultation. All the rating was done using seconds, results are presented as minutes and seconds; i.e. min:sec. We used intraclass correlation coefficient to test inter-rater reliability using SPSS version 14.0 Reliability Analysis intraclass correlation program.

Results

Clinical data entry

All four computer systems (EMIS LV, EMIS PCS, IPS Vision, iSoft) had similar methods for coding data even though they look quite different [9]. In all the systems problem titles are either coded from a picking list, or reselected if an existing problem. They also all had some type of standard form to speed up coding of chronic disease management; although they had different names and slight differences in functionality e.g. Templates and protocols in EMIS; SOPHIES in the iSoft system. These forms prompted the clinician to record all the clinically relevant data. All the systems also had prompting screens identifying the data items missing from patients' records and needed for the quality targets. The systems also allow free-text to be recorded. The principal differences between the computer systems were in appearance and the density of text. EMIS LV had a single screen through which data was entered and had the fewest lines of text while IPS Vision has the most with five windows providing access to the clinical record [10].

General observations about the consultations

The four hypertension monitoring consultations took between 7 and 11 minutes; blood pressure measurement was common to all the consultations and issuing of a repeat prescription for antihypertensive medication took place in all but one. In one consultation no items were coded the only use of the computer was to issue a repeat prescription. All the other clinicians coded between four and seven items in the consultation: blood pressure measurements, problem title or diagnosis and smoking status were the commonest and all form part of the quality target data requirements. The times taken on each aspect of the consultation are shown in Table 1. No referrals were made

Table 1 – Rating consultations: four simulated consultations using EMIS PCS, EMIS LV, iSoft and Vision

Consultation characteristic	Key **	EMIS PCS		EMIS LV (1)		EMIS LV (2)		IPS Vision		Summary	
		Time	%	Time	%	Time	%	Time	%	Median	IQR
General observations											
Duration		6:55		11:26		10:10		8:36		**9:13**	
Coded entries		0		5		7		4		**4.5**	
Time/ code		0		0:19		0:18		0:19		**0:19**	
First view: ***Interaction between actors (C+V = total verbal communication)***											
Computer-Dr	**Z**	0:03	0%	0:06	0%	0:11	0%	0:05	0%	**0:06**	0:03
Dr-computer	**X**	0:51	12%	2:42	24%	4:04	40%	2:24	28%	**2:30**	1:03
Dr-Patient	**C**	3:19	47%	3:36	32%	6:18	61%	3:36	41%	**3:36**	0:39
Patient-Dr	**V**	2:35	37%	4:30	40%	2:48	27%	2:24	29%	**3:42**	0:36
Second view: ***Data entry times (S +D = Total coding time)***											
Coding	**S**	0	0%	0	0%	0:49	10%	0:11	2%	**0:30**	0:19
QOF-code	**D**	0	0%	1:36	14%	1:18	13%	1:06	13%	**1:18**	0:18
Free-text	**F**	0	0%	1:06	9%	0:19	3%	0:41	8%	**0:41**	0:23
Prescribing	**G**	0:48	11%	0	0%	0:54	8%	0:36	7%	**0:48**	0:05
Third view: ***Body language in consultation (Q+T), Examination (E) and computer interference (W+R)***											
Eye contact	**Q**	4:18	61%	6:00	53%	4:30	43%	3:12	37%	**4:24**	0:54
Dr Comp&speak	**W**	0:29	7%	0:17	2%	2:18	22%	0:46	9%	**0:38**	0:40
Examination	**E**	1:12	17%	1:42	15%	0:58	9%	1:36	19%	**1:24**	0:30
Silent time	**T**	1:08	16%	2:00	17%	0:29	4%	0:58	11%	**1:00**	0:27
Pt-Dr&Comp	**R**	0:03	0%	0:18	2%	0:43	7%	0:38	7%	**0:28**	0:26
Reliability test: ***Intraclass correlation coefficient (ICC)***											
ICC Value		0.962		0.926		0.931		0.896			
95% CI		0.895 – 0.991		0.833 – 0.978		0.854 – 0.977		0.783 – 0.962			

Key: **Letters in column 2 are the ObsWin keys used to record elements of the consultation. Time = minutes: seconds
Z = Computer-doctor interaction; X = Doctor-computer; C = Doctor-patient; V= Patient-doctor interaction
S = Coding; D = Quality target coding; F = Free-text entry time; G = Prescribing time using the computer
Q= Eye contact; W = Doctor using computer and speaking; R = Patient speaking to doctor, while doctor uses computer;
95% CI = 95% Confidence Intervals.

(**A**) in any of the consultations and no new medications were started specifically to achieve the quality targets (**H**) so these lines are not shown in the table. The total times on each activity can add up to more than 100% as more than one activity maybe going on simultaneously.

ObsWin analysis

The majority of the consultation time spent on doctor patient interaction and communication. The patient speaking to the doctor varied between 2:42 minutes to 4:30 minutes (26-40%) of the consultation and the doctor speaking to the patient varied from 1:54 min to 6:18 min (32% to 61%). In three consultations the doctor spoke more than the patient and in one consultation the patient spoke more than the doctor. The doctor and patient spoke to each other for 70 to 88% of the consultation

Computer data entry was divided between entering coded data, free text and prescribing; referral was not needed in this consultation. The shortest and longest times using the computer were 51 seconds (12%) and 4:04 minutes (40%) using the computer. The other two consulters spent 2:24 minutes (27%) and 2:42 minutes (21%) on the computer. Coding time varied from 1:17 minutes to 2:06 minutes. It took between 18 to 19 seconds entering each coded item. Only three out of 16 items coded were for items not required for the hypertension quality target. The three of four doctors who made a free text record entered data for a mean of 42 seconds (median 41, range 19 to 66 seconds.) They recorded between one and four lines of free text, typing from four to 60 characters per line of text. Repeat prescription issue time was between 30secs and one minute.

All the GPs spent between six and eight and a half minutes on the consultation aside from computer use. There was no relationship between time spent on the computer and time taken to complete the consultation. The consulter who spent most time on the computer (40%) entered 7 coded entries into the computer and wrote one line of free text. This consulter still spent 5:54 minutes on the rest of the consultation. Another consulter spent 2:24 minutes on the computer (27%) coding four entries and typing three lines of free text. The rest of the consultation lasted 6:12 minutes. The last consulter who spent 2:42minutes (21%) of the consultation on the computer coded five entries and typed four lines of free text. This consulter spent 8:36minutes on the rest of the consultation. Finally the consulter who only used the computer 59 seconds for prescribing did not code any entries but spent 6:09 minutes on the rest of the consultation.

Not using the computer in the consultation appeared to allow more time for non-verbal communication and more computer use was associated with the need to communicate during computer use. Three to six minutes of the consultation time was spent on eye contact. Silent time was generally between 29 seconds to two minutes of the consultation. The consulter who spent the most time on the computer, also spent the most time speaking to the patient while using the computer, 2:18 minutes, whereas the other consulters only spoke to the patient 17 to 46 seconds of the time while using the computer. This consulter had proportionally the least amount of eye contact with the patient throughout the consultation (37%), whereas the consulter who spent the least time on the computer had the most eye contact during the consultation (61%). Patients speaking to the doctor while the doctor used the computer varied; but generally if the GP used the computer more the patient spoke to them more during computer use.

Examination time, in this case taking the patient's blood pressure, lasted 58 seconds to 1:42 minutes (9-17% of the consultation).

The type of computer system used did not seem to make a difference in the time spent entering coded data, the time spent entering coded data was very similar. Although their interface with the clinician varies greatly, this appeared to make little difference to the clinician using them who knew how to navigate to the appropriate parts of the program.

Reliability

We found good inter-rater reliability using intraclass correlation coefficient. The intraclass correlation coefficient of the consultations based on four or five observers ranged from 0.962 (95% CI 0.895-0.991) to 0.896 (95% CI 0.854 – 0.977). All these observations are significant at the ($p<0.001$). Most of the raters assessed two or three consultations. These estimates suggest that this method has a high inter-rater reliability.

Discussion

This technique to demonstrate how the computer now occupies around a quarter of consulting time, that most of our consulting GPs are coding 4 to 5 items and that these are predominantly data items required for the relevant quality target. We found that patient doctor verbal communication still remains the majority of time spent on the consultation, although the total time spent on the computer varied considerably between consulters. The type of computer system used did not have an impact on how much time was spent on clinical or quality target coding, clinician factors appeared to be much more important.

We have successfully piloted a novel but reliable tool for the analysis of the clinical consultation. We have demonstrated that it is feasible to combine: a carefully developed training package; multi-channel video and the use of ObsWin to achieve reliable results largely using student observers. Although it took at least two hours to analyze one consultation this is much less time consuming and more reliable than using a stop-watch and on screen timers to calculate the time taken as we have used in previous investigations [3,4]. Clinicians who participated in this experiment found the detailed observation of their behavior provided valuable feedback.

If this method shows similar reliability when used by other teams, it could provide a method for comparing the impact of quality targets on the clinical consultation as well as a mechanism for comparing different GP computer systems. It would allow the time taken to retrieve information, time

taken in data entry and the extent to which the computer appears to interfere with the consulting process to be analyzed. With larger samples, statistically significant differences might be apparent between consultations, which did and did not include quality target data collection and varying computer systems, both overall and in relation to other specific elements of the consultation (e.g. time taken to prescribe or refer.) It is possible that structured feedback might have a positive influence on consultation outcomes.

There are research methods and technical limitations to this study. This is a pilot study using simulated consultations with only a small sample size and a limited number of raters. Hypertension is a common disease that GPs frequently encounter and the 18-19 second consultation time reflects their ability to manage it. The consultations lasted from 7 to 11 minutes. UK GPs generally have 10 minute booked consultations and the consulting GPs said they felt that these consultations reflected their usual practice. Showing reliability with a small number of consultations and student raters is both a strength and weakness of this study. The method should be reliable in a larger sample, but has not as yet been tested in a wider range of consultations. If the tool is used in a wider range of consultations it may require clinicians to rate the consultations as it may not be possible to train non-experts in how to recognize coding in a wide range of clinical contexts. Some variables used in the consultation were easy to measure while others were more difficult. Use of keyboard and speaking time were the easiest except for very brief episodes. The raters found inferring whether there was eye contact and whether the GP was being provided information by the computer the most difficult elements of the consultation to assess. We did not test two of the items in the scale as the relevant activity did not take place in the consultation. The rating scale might be improved by adding a specific item for use of on-line information items or decision support.

The coding time in our study (18-19 seconds) is consistent with that reported for using Read Codes. This is consistent with the one previous UK study, though about half the time reported from a study from USA. The UK study measured the average time spent on coding as 14-27 seconds when using Clinical Terms and 18-49 seconds using Read Codes [11]. Cimino et al. studied the coding of data in an outpatient setting using videotaping as well and found that data entry time averaged 40.4 seconds per item [12]. Our results maybe shorter because hypertension is a very common chronic disease which GPs are faced with on a regular basis and the GP computer systems provide special data entry forms for this condition.

Further research is needed to see if this technique is reliable in other clinical contexts – both different clinical consultations and other specialties to compare whether consultations which involve quality target data recording are different from those without and to explore differences between clinical computer systems.

Conclusion

This novel multi-channel video technique is a reliable method for measuring patient – doctor - computer interaction; time spent on clinical and quality target coding; and other activities during the GP consultation. Computer use varied widely depending on the consulter's computing style, but the average time spent on the computer in these simulated hypertension consultations was 25%. Whilst entering quality target data and clinical coding took 18-19 seconds for each coded data item and 4 to 5 items were entered per consultation, patient - doctor communication, including verbal and non-verbal communication, still forms the overwhelming majority of the GP consultation time.

Acknowledgments

Informatics students for rating the consultations; Nigel Hague for clinical coding advice; Toby Fallon and Eben Wilson for technical help and Vickie Wood for reviewing the manuscript.

References

[1] Roland M. Linking physicians pay to quality of care a major experiment in the UK. N Engl J Med 2004; 351(14): 1448-54.

[2] Coleman T. Using video-recorded consultations for research in primary care: advantages and limitations. Fam Pract. 2000;17(5):422-7.

[3] Theadom A, de Lusignan S, Wilson E, Chan C. Using three-channel video to evaluate the impact of the use of the computer on the patient-centredness of the general practice consultation. Informatics in Primary Care 2003 11: 149-56

[4] Sheeler I, Koczan P, Wallage W, de Lusignan S. Low cost method for three-channel video assessment of the consultation. Informatics in Primary Care *(in press)* 2006

[5] Leong A, Koczan P, de Lusignan S, Sheeler I. A framework for comparing video methods used to assess the clinical consultation: a qualitative study. Med Inform Internet Med. 2006;31(4):255-65.

[6] Martin, N., Oliver, C., & Hall, S. ObsWin: Observational Data Collection & Analysis for Windows. London: Antam 2003.

[7] Byrne PS, Long BEL. Doctors Talking to Patients. Exeter: The Royal College of General Practitioners, 1984.

[8] Primary Care Informatics, St George's. Instruction Manual – How to measure clinical coding during the GP consultation. http://www.gpinformatics.org/reference/GPcoding.htm

[9] Primary Care Informatics, St George's. Recognising structured data and Screenshots of EMIS PCS, EMIS LV, iSoft, IPS Vision. http://www.gpinformatics.org/reference/data_recording.htm

[10] Jordan K, Porcheret M, Croft P. Quality coding in general practice computerized medical records: a systematic review. Family Practice 2004;4:396-412.

[11] Brown P, Warmington V, Laurence M, Prevost AT. Randomised crossover trial comparing the performance of Clinical Terms Version 3 and Read Codes 5 byte set coding schemes in general practice. BMJ 2003; 326:1127

[12] Cimino JJ, Patel VL, Kushniruk AW. Studying the human-computer-terminology interface. J Am Med Inform Assoc 2001;8:163-73.

Address for correspondence

Simon de Lusignan, Community Health Sciences, Hunter Wing, St. George's - University of London, London, SW17 0RE, UK. slusigna@sgul.ac.uk

Chapter 8.

Sustainability

This page intentionally left blank

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Sustainable Health Systems: Addressing Three Key Areas

Prajesh N. Chhanabhai, Alec Holt, George Benwell

Department of Information Science, University of Otago, New Zealand

Abstract

In the modern context sustainable health systems are being developed using the newest technological and communication technologies. This is proving to be a great success for the growth of Health Informatics and healthcare improvement. However this revolution is not being reached by a lot of the world population. This paper will address the importance of closing the Digital Divide, Empowerment of health consumers and the importance of converging communications. Key areas in the development of a truly sustainable health system

Keywords:

digital divide, empowerment, internet use, developing countries., communication technologies

Introduction

Health and Information Technology is an area that is recognised globally as being the next technological revolution. This is being helped by new technologies that have been devised to facilitate this transition e.g. Healthphone [1]. However this growth is only being appreciated by those that can access, afford and use it. There is a large population group that this Technological revolution is still leaving out. It is imperative to understand these before a country can build any truly sustainable health system.

This paper will highlight the issues that are faced by developing countries in Africa and may indicate that simply introducing an electronic medium will not solve problems related to the dissemination of health information. This paper advocates introducing technology but also paying attention to its effect and the importance in understanding cultural paradigms when it is introduced.

Background

The use of technology is growing in all areas of health communication, including consumer, and healthcare provider education, interaction as well as in the areas of decision and social support, health promotion, knowledge transfer and the delivery of services. This has lead to the development of a number of new fields such as captology the study of computers as persuasive tools and eHealth the integration of health care delivery and information delivery through computer based technologies[2]. The development of such areas is an indication of the immense potential for health communication efforts on a global scale [2].

Despite this development, the adaptation and integration of information technology in the health sector is unfolding at a slower rate than their counterparts in the finance and commerce sectors [2]. Currently most developments that include the health consumer are run by for-profit eHealth companies. These companies currently utilise the Internet and web related technologies to run their organisations. The most common focus of these organisations is to provide tools, solutions, products or services that aid some aspect of clinical care or eCommerce. However, as most of these organisations are for-profit, their impact on areas and communities that are economically challenged is minimal or non-existent. As it currently stands these facilities can only be accessed by those that have access to the Internet and the ability to pay for such services, hence segregating a large number of the population in many countries and in a majority of the cases this is the population that needs the greatest need for eHealth services [3].

According to Kreps many of the people who are at most risk from serious health conditions come from underserved populations, populations that are generally made up of individuals who are of low socioeconomic status, possess low level of health literacy and are members of marginalised ethnic and minority groups [4]. These underserved and vulnerable populations often have limited access to relevant health information especially information that is otherwise easily available over the Internet. This is one of the symptoms of the Digital Divide, however within the health sector the Digital Divide as a more specialised problem. Many of the characteristics that identify those on the have not side of the Digital Divide also apply to those who suffer from the negative effects of health disparities. While information and knowledge are not guarantors of good health care decisions and adherence to recommended health behaviour, their ease of availability has shown to contribute to them [5][6]. This has been recognised by the White House, who in their *Healthy People 2010* report indicated that health communication through the use of computer technologies is a means of bridging the digital health divide [7].

Table 1 - World Internet Usage and Population Statistics. Highlighting the impact of the Digital Divide in Africa. Adapted from Internet Usage Statistics The Big Picture [9]

World Regions	Population (2006 Est.)	Population % of World	Internet Usage, Latest Data	% Population (Penetration)	Usage% of World
Asia	3,667,774,066	56.4 %	394,872,213	10.8 %	36.4 %
Europe	807,289,020	12.4 %	308,712,903	38.2 %	28.4 %
North America	331,473,276	5.1 %	229,138,706	69.1 %	21.1 %
Latin America/Caribbean	553,908,632	8.5 %	83,368,209	15.1 %	7.7 %
Africa	915,210,928	14.1 %	32,765,700	3.6 %	3.0 %
Middle East	190,084,161	2.9 %	19,028,400	10.0 %	1.8 %
Oceania / Australia	33,956,977	0.5 %	18,364,772	54.1 %	1.7 %

Digital divide

Information is critical for the development of any system. This is of crucial importance in health systems, in which the transfer of correct information maybe the difference between curing a patient and killing a patient. Information flow can be regarded as one of the most important factors for improving health systems especially in under-developed resource poor countries. The importance is even greater than worrying about the availability of infrastructure, the increasing of health workers and the distribution of funds, because without the correct information flow none of these other factors can function effectively or efficiently, and in most cases even in fthe correct settings. Access to information is essential in developing countries as it allows health policy makers in these countries to understand their current deficiencies thus enabling them to develop practical ways in which to solve these deficiencies. Information can thus be seen as a form of empowerment however this empowerment can only be useful if it comes from correct information, this is where the dissemination of information is important [8].

Advances in information and communication technologies have inherently made the distribution of health information globally seem effortless. The technologies in particular the Internet, allows information to be made available the instant it is produced. The information that is transmitted by these technologies can allow users to choose the exact type of information they want and can view the information in a number of different mediums. This information can then be accessed by a multiple range of users, be it a farmer in rural Africa or a health policy government official. However, this is a utopian idea as it is rare for a farmer in rural Africa to have access to the Internet.

The information gap between the developed world and developing countries is currently widening, this is aided by the growth of the Digital Divide. The impact of the Digital Divide is so large that the [10][8][11] have all commented that it is more dramatic than any other inequity in health or income. This fact is frightening as the development of information and communication technologies was prophesized has being one of the solutions for these other inequities. The extent of the Digital divide can be gathered from the following statistics. Africa has a population of over 900 million which makes up 14.1% of the world population. Only 32 million of the entire population have access to the Internet, with a penetration percentage of 3.6%. Egypt, Morocco, Nigeria and South Africa make up 60% of the total Internet users in Africa. Table 1 shows the global Internet use by region. The table highlights that despite having the second largest percentage of the worlds population, the Internet use in Africa can be regarded as being the lowest per head. According to the United Nations Development Program There are more Internet hosts in New York than in continental Africa more hosts in Finland than in Latin America and the Caribbean [12]

These statistics show that the digital divide is growing larger, and in the health context this is contrary to the aim of medical research that is being produced. There is now numerous research into how the Internet and such pervasive communication technologies will play an important role in improving health outcomes. There is also the concern that if this technology is introduced or made available in developing countries that it will not be used [13]. Norman and Skinner have reported in their paper that in the United States of America and Canada alone over 40% of adults have low basic literacy levels, thus e-health resources are likely to be inaccessible to large segments of the population [14].

The question that then begs to be answered is if the technology is available in such countries and is not being used then how effective will it be in places like Africa? Using information technology does not mean just placing the infrastructure and expecting the health consumer to buy into it. Using Information technology in the ehealth context requires the user to have a certain level of e-health

literacy, the ability to read, use computers, search for information and to understand the information in context [14]. In addition to these fundamental factors there is also the need to address the cultural dimensions of change. In developing countries like Africa, more than any other place, there are very strong ties to the traditional beliefs and these need to be factored into the development of a good health information dissemination system.

Empowerment

Traditionally the healthcare consumer has been the least consumer like and the least informed [15]. Protection from social stigma as well as the feeling that patients would get more sick once they knew their medical condition were the reasons used by many traditional physicians to discourage empowerment With expanding populations and the increase in occurrences of epidemics, medicine has become more scientific and thus medical knowledge has started to become available to the lay public [15]. In the last twenty years the emphasis has changed from cure of health conditions to prevention, with an emphasis on health and wellness [16]. According to Amatayakul *Patients have become interested in making **choices** for themselves about their physicians, treatments and lifestyles* [15], this can be clearly observed in the change of terminology from medical care to healthcare. The term medical care focused primarily on processes administered by a physician, whereas healthcare encompasses a broader range of services and procedures [17].

The emergence of communication technologies and the incentives in the health sector to include consumers in their operations are some of the factors in increasing the importance of the consumer in the healthcare setting [17]. The biggest factor in cementing the role of the patient as a consumer is through the growth and the innovative capabilities of technology. The increasing availability of interactive information has enabled many services to be made available online.

boards and bulletin boards has allowed individuals to share experiences of specific diseases and treatments. This has introduced another dimension to the healthcare industry where consumers are more knowledgeable and understanding of the terminology and procedures that are used in the health sector.

According to Eysenbach, initially the technology had been looking at development and growth through the eyes of the medical professional, with the drive towards consumerism[18]. This has changed and has seen the birth of consumer health informatics. Consumer health informatics is defined by Eysenbach as *the branch of medical informatics that analyses consumers' needs for information; studies and implements methods of making information accessible to consumers; and models and integrates consumers' preferences into medical information systems*. [18]. This definition agrees with Amatayakuls statement that the *principle of consumer health informatics is that of empowering individuals to play a greater role in their own healthcare and to be active participants in the decisions that affect their healthcare* [15]. Figure 1 shows how consumer health informatics has shifted from individuals just having a choice to being empowered. Essentially, all health information can now be used to benefit the consumer of healthcare.

Empowerment will only occur if the consumers themselves are allowed to interact with the healthcare system. An interaction in which they do not receive just limited feedback, but one that promotes two way feedback, benefiting all parties. This idea has been adopted by the European Union, who realised that the greatest interaction would occur if patients had access to their records [18]. In October 1998 the European Union required that each of their member countries passed legislation that would ensure consumers in those countries to have access to all their health records [18].

In the developing world context, this empowerment will be a much harder process to implement due to the nature of lifestyles and beliefs in such countries. As was mentioned earlier it is the shift in cultural paradigms that needs to be addressed. Empowerment does not have to occur with only the use of technology. The important factor is getting the information out in a manner that is both timely and comprehensible to the lay public.

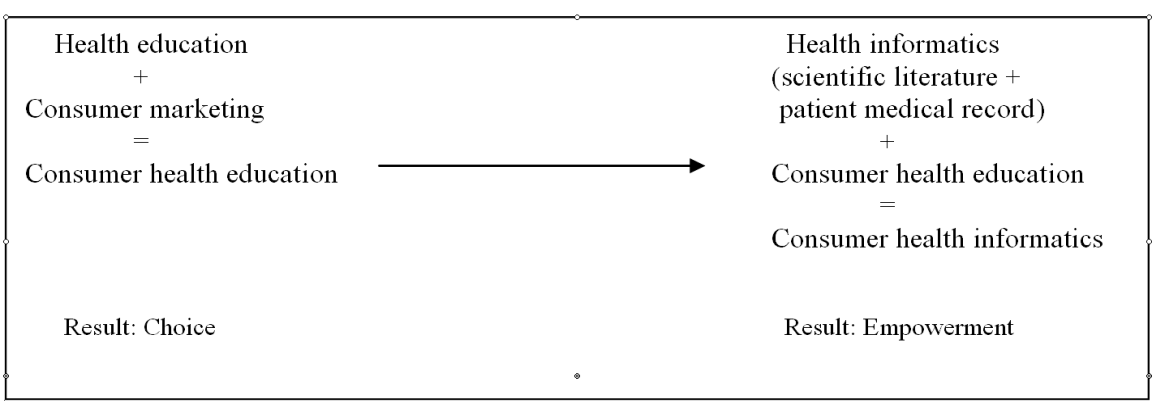


Figure 1 - The shift from an individual having a choice to being empowered. Adapted from Amatayakul [15]

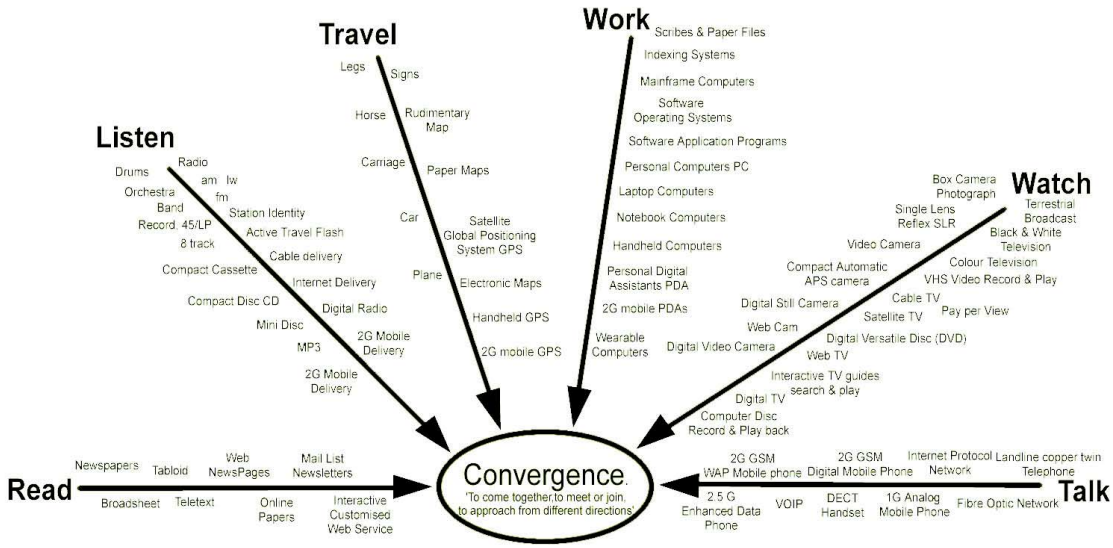


Figure 2 - Diagrammatic representation of the convergence of communications

A good example of this is the difference between the HIV/AIDS policy in Uganda and Botswana. Both countries are known to have some of the higher HIV/AIDS infection rates in the world. Both countries established different schemes to bring down the infection rates. The success rates in both countries have differed considerably.

The main documented reason was due to the condom promotion. According to Allen and Heald, in Botswana condom promotion provoked antipathy from church groups, local healers, parents and chiefs. These all being very focused on the cultural shift and the failure in understanding why this was being done, they had not been informed correctly as why the process was occurring. Conversely in Uganda condoms were not initially introduced, rather the presidents role in being vocal about the problem had an effect of inspiring local councillors, chiefs and church leaders to take this as a directive to educate their people about changing their sexual behaviour [19]. This program is now being attempted in Botswana. This example indicates the importance of disseminating health information via the right channels to obtain the best possible result that was aimed for.

Empowerment is feared by many in the medical field as care givers feel that this knowledge can harm rather than hurt the patient. However, in this case empowerment is through education. With a global focus on the Internet and its capabilities, researchers must not forget those areas that cannot harness these technologies. Rather, it is imperative to find a way that empower, through education via the correct channels

Converging communications

It is thus imperative to examining technology based health communication to understand who and what the communication is designed for. The goals of health promotion and disease prevention communication efforts are to help health consumers and information seekers gain knowledge about health issues and improve health. The goals of communication of health care delivery are to treat illness, maintain or improve health among patients and increase cost and delivery efficiencies[23]. Health communication efforts are designed to improve lifestyle behaviours, reduce risk factors for disease, increase compliance with a medication or treatment plan, better self manage a condition, provide social support or provide help with decision making procedures[24].

Through new technologies, health promotion and disease prevention interventions are being delivered successfully on-line, on CD-Rom, over the telephone, through handheld computers (PDAs), and via other technologies (Figure 2) for a variety of topics including weight control, injury prevention, smoking cessation, nutrition promotion, and medication compliance [20][21][22]. New technologies also allow health care delivery to transfer its model of care into a model of telemedicine, consisting of, but not limited to, telephone, video, and e-mail consultations, e-prescribing, claims processing, physician Web portals, and Electronic Health Records [23][2][24]

Figure 2 describes emerging technologies and trends that are singularly powerful. Their convergence could shift basic paradigms in health and health care. Potential examples of such converging applications include wireless, subcellular biosensors that monitor individual health parameters in real-time; techniques for meta-analyses of genetic, biophysical, and behavioural information to inform development of personalised health interventions including therapies; and tailored, broadband, interactive multimedia health communications that occur, irrespective of economic background[24].

The merging of consumer informatics and health communication have a combined effect of focusing on how communication methods will have an impact on consumer decisions. Consumer informatics aims to shift public knowledge, motivations, and attitudes towards clinical behaviours and with the adoption of health communication strategies this will yield more interactive, flexible and multidimensional healthcare tools.

Conclusion

This paper has highlighted the fact that Information technology is important in the health sector. Its role is growing in leaps and bounds. However, its growth will be hampered by its lack of accessibility in key areas. These areas lie in the developing world, primarily in Africa. It is also these areas that have the lowest Internet access rates and the highest rates of deaths by infectious diseases [25]. The Internet may not be the solution in dissemating health information in economic and poverty stricken areas, rather some of its methodologies can be implemented in other ways to provide optimal solutions.

In order to build sustainable health systems, it is important to understand why current systems will not work in specific areas, and to be able to adapt these systems. These adapted systems will then be more productive and effective in the areas that need it the most. One way of achieving this goal is to look at the converging of various technologies. This will include both electronic and non-electronic means to achieve the communication framework that will allow the right health information to be disseminated to the right people.

References

[1] Cuzens S. The power of cooperation. Bright. 2006:17:28- 31.

[2] Suggs SL. A 10-year retrospective of research in new technologies for health communication. Journal of Health Communication.2006:11: 61-77.

[3] Eng TR. The ehealth landscape: a terrain map of emerging information and communication technologies in health and health care. The Informatics Review. 2001 [Online] Available http://www.informatics-review.com/thoughts/rwjf.html

[4] Kreps GL. Disseminating relevant health information to underserved audiences: implications of the Digital Divide pilot projects. Journal of the Medical Library Association. 2001: 93: 68-73.

[5] Tang PC. and Lansky D. The missing link: bridging the patient-provider health information gap. Health Tracking.2005: 24: 290-1295.

[6] Harris L, Dresser C and Kreps GL. E-Health as dialogue: communication and quality of cancer care. Dialogue Systems for Health Communication. 2004 [Online] Available http://www.ccs.neu.edu/home/bickmore/dshc/kreps.pdf

[7] Department of Health and Human Services. Healthy People 2010. 2000 [Online] Available http://www.cdc.gov/nchs/about/otheract/hpdata2010/abouthp.htm

[8] Godlee F, Horton R and Smith R. Global Information Flow. Biomed Central News and Views. 2000:1:4-6.

[9] Internet Usage Statistics- The Big Picture. 2006 [Online] Available http://www.internetworldstats.com/stats.htm

[10] [10] Fors M. What the United Nations can do closing the Digital Divide. UN Chronicle Online Edition. 2005 [Online] Available http://www.un.org/Pubs/chronicle/2003/issue4/0403p31.asp

[11] Edejer TTT. Disseminating health information to developing countries: the role of the Internet. BMJ. 2000:321:797-800.

[12] United Nations Development Programme. Report of the Meeting of the high-level panel of experts on information and communication technology. 2000 [Online] Available www.un.org/documents/ecosoc/docs/2000/e2000-55.pdf

[13] Jensen M. African Internet connectivity. 2006 [Online] Available http://www3.sn.apc.org/africa/

[14] Norman CD and Skinner HA. eHEALS: the ehealth Literacy Scale. J Med Internet Res 2006:8:(4):e27. [Online] Available http://www. Jmir. Org/2006/4/e27/

[15] Amatayakul M. EHRs and the consumer: a new opportunity. In Murphy GF, Hanken MA and Waters KA eds. Electronic Health Records Changing the Vision. Philadelphia: WB Saunders Company: 1999: 26-68.

[16] Mongerson P. A patients perspective of medical informatics. Journal of American Medical Informatics Association. 1995:2 (2): 79-84.

[17] Friedman M. How to cure health care. Public Interest. 2001:142 (Winter).3-30.

[18] Eysenbach G. Recent advances: Consumer health informatics. BMJ.2000:320. 1713-16.

[19] Allen T and Heald S. HIV/AIDS policy in Africa: What has worked in Uganda and what has failed in Botswana? J. Int. Dev. 2004:16:1141-1154.

[20] Anderson DG. and Stenzel C. Internet patient care applications in ambulatory care. Journal of Ambulatory Care Management. 2001: 24:138.

[21] Dijkstra A, De Vries H, and Roijackers J. Computerized tailored feedback to change cognitive determinants of smoking: A Dutch field experiment. Health Education Research. 1998:13:197206.

[22] Friedman RH, Kazis, LE., Jette A, Smith MB., Stollerman J, Torgerson J, and Carey K. A telecommunications system for monitoring and counselling patients with hypertension. Impact on medication adherence and blood pressure control. American Journal of Hypertension. 1996: 9:285292.

[23] Chhanabhai P. EHRs: Fear of breach? The New Zealand publics opinion. Unpublished Masters thesis. University of Otago. 2006.

[24] HIFA2015. Communications with members of the HIFA2015 forum [Online] Available http://www.dgroups.org/groups/hifa2015/

[25] World Health Organisation. World Health Statistics 2006. 2006 [Online] Available http://www.who.int/whosis/en/index.html

Address for correspondence

Prajesh Chhanabhai
Department of Information Science
University of Otago
PO Box 56
Dunedin
New Zealand
Email: pchhanabhai@infoscience.otago.ac.nz

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Locating Nursing Classification Schemes within Health Information Strategies for New Zealand

Shona K. Wilson[a], Jan Duke[b],

[a] *PhD Student, Graduate School of Nursing, Midwifery and Health, Victoria University of Wellington, New Zealand*
[b] *Professor of Nursing and Head ,Graduate School of Nursing, Midwifery and Health, Victoria University of Wellington, New Zealand*

Abstract

The potential to use classification schemes to describe and measure nursing in a country that has previously not used them as a part of practice is fraught with issues. Such is the case for New Zealand. Without nursing specific classification scheme use in the information systems of day to day function, nursing cannot validate what it does and the difference it makes to health outcomes for New Zealanders.

The local use of valid and reliable classification schemes as tools to capture locally generated data that is able to be used as quality data needs to be considered alongside the national use of reliable clinical reference tools that are consistent with international standards. This may make the difference to the potential for significant contribution of nursing practice specific data to health information collections in preference to a 'one fits all' approach to user interface nursing classification scheme adoption at a local level. Tensions between a top-down approach and a locally based bottoms-up practice based approach and associated issues provide the core to this paper.

Keywords:

nursing classification schemes, national health information strategies, New Zealand

Introduction

There seems to be little doubt in the international world of informatics that the collection of nursing specific data can contribute to the measurement and validation of nursing practice [1]. However how to ensure that data can ultimately contribute to national and international health knowledge bases in a manner that will be sustainable over time remains emergent. The fabric of information collection, storage, retrieval, communication for optimising use is complex and involves many layers of key stakeholders ranging from users, developers, administrators and educators through to decision and policy makers. Unless the data collected is high quality data as defined by Kerr, (2006) and fits within provider based generic and national information systems frameworks, the information and knowledge it will ultimately provide will be compromised [2]. A tension arises between the nature and frequency of the data collected at the user interface and the multiple uses for each piece of datum within the health sector through to a policy level.

Contemporary information systems use by nurses in New Zealand appears to be largely focussed on information retrieval rather than data input. This means that although the retrieved information may contribute to nursing decision making it will not represent the actions and outcomes that are considered to be the basic elements of nursing practice. Therefore for nursing the discipline specific data-information-knowledge continuum remains unachievable. For nursing specific data to be useful it must represent nursing practice within clinical information systems in a way that is internationally, nationally and locally comparable [3] and yet satisfy the perceived need within the work force. Despite legislative and professional need for clear, concise, timely, accurate and current client record keeping by nurses processes remain problematical within the profession. Most documentation is as hardcopy files despite the presence of information systems in practice. Alternatively some free text descriptions exist within systems that do not contribute to wider knowledge repositories. The problem with free text is that it cannot be readily classified within the context of quality classification of health related phenomena.

At the same time there seems to be recognition of the need for measurement of health outcomes [4]. This tends to result in the application of a variety of tools being created for data collection at a local level. Many of the tools do not appear to have a theoretical or research base and most if not all do not appear to be integrated into clinically based information systems. The result is that existing clinical information systems do not and cannot meet the basic needs of nursing practice and consequently cannot contribute to the health information collections at a national level. To do this nursing must begin to collect the activities and outcomes of nursing practice using a reliable and validated discipline specific classification scheme that is integrated within the clinically based electronic client record [3]. In turn the electronic client record must contribute in some way to wider health information collections.

One issue for New Zealand is that clinical information systems are not yet in widespread use amongst major health providers although Patient Management Systems (PMS)

are. The next generation in systems development for the collection of clinically based data, the clinical information system (CIS), will provide the avenue for health professionals, including nurses to capture their daily activities and generate aggregate health outcome measurement [3]. Success of this depends upon a high level of participation at the user level to input data at the point where care is delivered.

Nurses appear to not know about or understand the differences between management/administrative systems and clinically based information systems. Nor do they appear to discern clearly between medical and nursing models of care delivery. Concurrently clinically based systems development seems to rely upon clinician input. When this happens nurses working at the clinical face tend to opt for transposition of what they already know into the technological world, often resulting in the mechanisation of the familiar paper based world. As classification schemes for nursing are not part of the paper nursing world in New Zealand subsequent implementation is unlikely to be driven by clinicians. Success is dependent upon nurse users in large numbers accurately recording the activities of practice to produce quality data, reliable information and knowledge. To do this they must understanding the meaning of each user interface category.

A solution may be to use a nursing classification scheme at the user interface that has high reliability and validity test scores. A problem is that such classification schemes have been developed and tested outside New Zealand and may not necessarily fit the nursing context of the country outside the country of origin [5]. Every change that occurs at the user interface has the potential to reduce the reliability of the tool and subsequently the quality of the data collected and the potential to be useful beyond the local context.

National trends in health information collection

Each country while acknowledging the place in and influences of the global information community has differences within the way health care is funded and provided that impact upon potential for quality data collection within information systems. The more competitive commercial approach of systems development may well carry with it a trade-off in quality of data collection than the more collaborative approach of a top down nationally supported model. Clearly for successful implementation of nursing classification schemes to collect nursing specific quality data within clinically based information systems an approach is needed that enables local, regional and national collections. The collections must be able to compare like data and make relationships between that data to validate the activities and outcomes for nursing in the context of health information collections. Then the difference that nursing makes to the health outcomes for New Zealanders will be able to be measured.

New Zealand as a country with a small population and a single government has some advantages over larger more complex societies for health information technology implementation. Health information strategy and implementation is a government direction resulting in the widespread public consultation and uptake of the Working to Add Value through E-information (WAVE) (2000) [6] document and more latterly the New Zealand Health Information Strategy (HIS-NZ), (2005) [7]. The government led approach to national health information collection has to date resulted in the implementation of a National Health Index (NHI), which categorises the demographic information about each member of the population, who accesses the health care system. Each person is subsequently registered with a alpha-numerical unique identifier that is used by all providers of health care and disability services. The NHI identifier and associated information is used to help with the planning, co-ordination and provision of related services across New Zealand. It is also intended to improve the flow and sharing of health information across providers and locations. The Medical Warnings System (MWS), associated with the NHI, warns healthcare providers of any known risk factors that may be significant when making decisions about care. The providers with electronic access can access the national NHI and MWS repository and automatically populate their PMS accordingly. The NHI and MWS do not contain any clinical information. A National Minimum Data Set (NMDS) for hospital events is available from a national level for providers of hospital services. No NMDS has yet been developed for use by primary and tertiary providers. Consumers of health care can request and receive any information about them that is stored within any of the databases that are held at a national level associated with the NHI [8].

A current development toward the national collection of health information include the establishment of a national Health Provider Index (HPI) as a register of health providers and organisations to enable role based secure electronic access to national repositories [8] Concurrently there is a public consultative process taking place to establish key direction strategy for information collection in the primary health care sector.

Together the NHI, MWS and HPI provide the basic elements of the international development of National Nursing Minimum Data Sets (NNMDS) [9] except for the inclusion of standardised nursing terminologies. The challenge will be to find a find a way to integrate these elements with nursing classification schemes implemented at a local level.

PMS are now established within the secondary care sector, which is predominantly provided by publicly funded District Health Boards (DHBs). Many of the PMS are supported by CIS, which may provide the avenue for clinician based field data input.

However the concept of an Electronic Health Record (EHR) that captures longitudinal lifespan health details of the population and that is accessible regardless of location or provider remains the ‘holy grail’ of health data collection and flow for New Zealand [10]. This may be concerned with two significant factors. The primary factor is that the supply of information systems for the health sector is driven by

commercial companies. In tandem, standards and recommendations from the government are intended as a guideline and are not mandatory. All the government can do to encourage vendors to adhere to any associated standards is to provide incentives for vendor buy-in.

The introduction of commerce in the health sector is a relatively new phenomenon to New Zealand. There was a highly regulated mostly publicly funded health sector until the mid-80's when a process of de-regulation and revision of the health care system occurred. This introduced a more business supply/demand model and choices to access private sector health care began. Since that time the presence of commerce in health care delivery has mushroomed and in particular to do with health information technology.

Health information systems vendors have valued clinical input to the planning and design phases of the Systems Development Life Cycle (SDLC) [11] for the systems developments which have resulted in some systems being developed that provide a framework for the system. This enables service providers to configure according to their own unique needs. In some instances this has resulted in a variety of developments within common systems with little or no interoperability between them. Local needs do not always concur with regional and national needs and local systems can adopt a life of their own.

Most significantly other than the NHI and MWS there is a lack of standardisation at the local level, so that when any EHR becomes a reality, the ability to compare like elements of data will be reduced.

The development and introduction of standards for the collection of meta-data at a national level [12] will hopefully be taken up and adopted by the process of government led public consultation. However that is likely to take some time, while aficionados of health care development continue to develop as a business means. This has an impact upon the potential to collect nursing specific data. Classification schemes for the collection of the basic elements of nursing practice exist internationally. Consistently they have been developed to be used at the clinical nurse user interface [13]. Terms and associated definitions are accordingly attuned to the user interface. If in this climate of health data collection movement any are integrated within existing systems at a local level without cognisance of the national movement the potential for descriptions and measurements for nursing beyond that level will be negated.

Nursing classification schemes

For any nursing classification scheme implemented in nursing to gather quality data in New Zealand the scheme must be simple to use, attract large numbers of nurses as users and suit the context of the service. Like most countries nursing practice in New Zealand is delivered throughout diverse settings alongside disparate health information systems. It is doubtful that a single reliable and valid scheme will suit all nursing settings [13]. This paper proposes that a variety of schemes can be implemented to suit each unique setting, whilst retaining quality of data. The solution is to implement known reliable and valid schemata that are an exact coding match to cross-map with multi-disciplinary clinical reference terminologies. This broadens the scope of choice of a classification scheme to implement and guarantees a return of quality data.

The issue with this is that to date New Zealand has not committed to a national clinical reference terminology (such as SNOMED CT). A solution would be to use the International Classification for Nursing Practice (ICNP) as a clinical reference terminology ensuring that the terminology used at a local level is an exact coding match. When a national multi-disciplinary clinical reference terminology is implemented it should then cross-map accordingly. Negotiations between SNOWMED CT and ICNP began in April [14] and are currently taking place to ensure that there will be a cross-map available in due course.

Currently some of the major nursing classification schemes that have been validated as reliable and useful by research in practice internationally are validated as cross-mapping to SNOMED CT directly, without the use of an intermediary nursing specific clinical reference terminology [15]. However without the use of a nursing specific clinical reference terminology in New Zealand the potential for nursing data to be compared across international nursing borders will be reduced. But can nursing in New Zealand wait until there is an exact cross-map between user interface languages, the discipline specific clinical reference terminology (ICNP) as outlined as International Standard (ISO) 18104 [16] and the multi-disciplinary clinical reference terminology? When that occurs, a small country like New Zealand must then be able to purchase the terminologies within the global market and implement in a resource scarce health care sector.

Compounding these issues is that nursing classification schemes as they exist elsewhere have not been adequately tested to see if they describe the basic elements of nursing practice in the context of New Zealand. Do the terms and definitions within the classification schemes accurately summarise the nuances and differences in practice between New Zealand and the country of the terminology's origin? Clearly the need to begin to collect the nursing specific data outweighs the temptation to develop a local classification scheme which would take time and risk compromising the potential to accord an exact cross-mapping with international reference terminologies. Analysis and evaluation as testing of any terminology in the local context would appear to be an imperative in the cycle of development and implementation as suggested within the SDLC [13]. Unfortunately historically it appears that this important evaluative step is often omitted in New Zealand.

Like other countries New Zealand lacks the nurses who are adequately prepared to begin this journey. Whilst there is a small core of nursing informatics enthusiasts the call for their expertise is shared between academic research and education, systems development, local clinical information needs and nationally led strategic approaches. Such personnel need in-depth understanding of the top down

approach, insight into the local nursing systems development approach, the ability to identify with international trends and location of the issues within the socio-cultural context for resource allocation in New Zealand.

Is there a solution?

The way forward for nursing in New Zealand to begin to collect discipline specific data that will ultimately describe and measure the basic elements of nursing practice needs to be carefully planned in consideration of these contextual issues. To date there is not much available theory driven nursing informatics research literature available to inform such an undertaking as the focus has been historically driven by problem solution and subsequent systems development [11]. The work of Effken (2003) provides a conceptual framework toward organising informatics research based upon the constructs that make up the meta-paradigm of nursing. The aim of her framework is to provide a means to develop nursing informatics theory development as a research endeavour. The work locates the SDLC within the major constructs of nursing practice and emphasises the need for ongoing analysis throughout the cycle. Effken (2003) invites nursing academics and researchers to apply the main principles of the framework to nursing informatics theory development internationally and so validate its utility. This may provide a theoretical grounding for the implementation of classification schemes set in the context of the body of nursing knowledge that is New Zealand nursing practice.

The time has come to move beyond the dualistic methods of having to choose between the top-down and bottom-up approaches. A more hybrid approach that considers and values the issues of each may provide the solution for nursing classification scheme implementation in New Zealand. Acknowledging that a universal nursing classification scheme is unlikely to fit the diverse settings in New Zealand, implementation can begin by selecting a nursing classification scheme that is:

- valid and reliable internationally for the collection of quality data
- an exact cross map with international clinical reference terminologies, and
- matches a particular nursing service practice setting in New Zealand

If all implementation at local level adheres to these three basic criteria for implementation then the data collected and stored will be able to be used in many ways. It will be able to be used at a local level for care planning, measurement of service delivery, and begin to describe the difference nursing makes to the local population targeted by the service. It will also ensure that ultimately the data can contribute to other repositories as they are developed, such as regional and national EHR repositories and any professional nursing database that may evolve in the future.

Different classification schemes can be used in different settings as long as they have the exact cross map at coding level of CIS with the clinical reference nursing specific and multi-disciplinary clinical reference terminologies that are in use internationally. Some classification schemes lend themselves more to a particular service than others. For instance the Omaha System appears to be a good match for any New Zealand Mental Health Nursing Service. Mental Health in New Zealand has Health of the Nation Outcome Scale (HoNOS) as mandatory for the collection of Mental Health data [17]. There appears to be a high level of synergy between the Omaha System and HoNOS in the identification of problems as well as the language and structure of outcome scales. While a match between the two has not yet been tested, this does mean that nurses would recognise the similarities and so may be better placed to adopt the Omaha System in practice. The Omaha System has also been evaluated and demonstrates a high level of utility for Mental Health [18]. Another classification scheme may have greater match in another area such as acute nursing service but as long both have an exact match with the overarching clinical reference terminology, the data collected will be able to be used across local, regional, national and international settings.

Conclusion

The central argument for this paper is that for nursing practice to be able to validate what difference it makes to the health of New Zealanders and contribute to repositories concerned with the body of nursing knowledge and development, New Zealand needs to consider the introduction of both the nursing specific clinical reference terminology (e.g. ICNP) and a multi-disciplinary clinical reference terminology (e.g. SNOMED CT) at a national level. These reference terminologies would need to integrate with the NHI, MWS and HPI, as part of national health information trends. There is also a need to ensure that any implementation within any system at a local level is an exact match to the clinical reference terminologies. In turn an exact match between coding of the system and terms and definition of the validated and reliable nursing classification scheme needs to be ensured. If this occurs, then in time all the data that is collected at the user interface will ultimately contribute to the national health information collections of the future as well as to the national and international body of nursing knowledge.

Acknowledgements

The Foundation of Science Research and Technology, New Zealand, for a scholarship to undertake this study

References

[1] Harris, M. R., Graves, J.R. Solbrig, H.R. Elkin, P.L. and Chute, G.C. Embedded Structures and Representations of Nursing Knowledge. J Amer Med Inform Assoc 2000

[2] Kerr, K. (2006) The institutionalisation of data quality in the New Zealand Health Sector. 2006 Unpublished thesis at University of Auckland, New Zealand.

[3] Androwich, I.M., Bickford, C.J., Button, P.S., Hunter, K.M., Murphy, J. & Sensmeier, J. (2003) Clinical Information Systems: A Framework for Reaching the Vision. Washington, D.C: American Nurses Publishing.

[4] New Zealand Health Information Service. (2006). About NZHIS. Online at: http://www.nzhis.govt.nz/aboutnzhis.html [Retrieved 02 October, 2006]
[5] Clark, J. (1998) The International Classification for Nursing Practice Project Online Journal of Issues in Nursing. Available at htttp://www.nursingworld.org/ojin/tpc7/tpc7_3.htm [Retrieved 27 July, 2006]
[6] Ministry of Health. From strategy to reality: the WAVE project. Report of the WAVE Advisory Board to the Director-General of Health. Wellington, New Zealand: Ministry of Health; 2001.
[7] Health Information Strategy Steering Committee. Health Information Strategy for New Zealand, 2005. Online at: http://www.moh.govt.nz/moh.nsf/0/1912064EEFEC8EBCCC2570430003DAD1/$File/health-information-strategy.pdf [Retrieved 06 June, 2006)
[8] New Zealand Health Information Service. National Health Index, nd. Online at: http://www.nzhis.govt.nz/nhi/brochure.html#08 [Retrieved 20 June, 2006]
[9] MIA-NI. Nursing Matters: International Nursing Minimum Data Set (i-NMDS), 2004. Online at http://www.icn.ch/matters_i-NMDS_print.htm Retrieved [06 June, 2005)
[10] Kerr K. The electronic health record in New Zealand. Health Care and Informatics Online. 2004 Online at: http://hcro.enigma.co.nz/website/index.cfm?fuseaction=articledisplay&FeatureID=040304. [Retrieved 25 July, 2005].
[11] Effken, J.A., (2003) An Organizing Framework for Nursing Informatics Research in Computers in Nursing Vol.21 No. 6 pp 316-325. Nov/Dec 2003
[12] [12] New Zealand Ministry of Health: Health Information Standards Organisation (HISO) Online at: http://www.moh.govt.nz/moh.nsf/f872666357c511eb4c25666d000c8888/9a18e48fca47e93dcc256e2200071d6d?OpenDocument [Retrieved 26 November, 2006]
[13] McCormick KA., Jones CB. Is one taxonomy needed for health care vocabularies and classifications? Online Journal of Issues in Nursing 30 September 1998. http://www.nursingworld.org/ojin/tpc7/tpc7_2.htm [Retrieved 06 April, 2005]
[14] SNOMED International: ICN and SNOMED Further Collaboration. 2006. Online at: http://www.snomed.org/news/documents/icn.pdf [Retrieved 20 November, 2006]
[15] SNOMED International Nursing Working Group n.d. Mapping. Online at: http://www.snomed.org/clinical/SNOMEDMapping.html [Retrieved 06 June, 2006]
[16] International Organisation for Standardisation. ISO 18104: 2003. Health Informatics – Integration of a Reference Terminology for Nursing. Online at: http://www.iso.org/iso/en/Catalogue DetailPage.Catalogue Detail?CSNUMBER=33309 [Retrieved 01 August, 2006]
[17] The National Centre of Mental Health Research and Workforce Development: Te Pou o Te Whakaaro Nui. HoNOS Family of Measures, n.d. Online at: http://www.tepou.co.nz/page/tepou_83.php Retrieved 06 October, 2006
[18] Barrera C, Machanga M, Connolly PM, & Yoder M. Nursing care makes a difference. Application of the Omaha System. Outcomes Management Journal, 2003 Oct-Dec;7(4):181-5.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

SHARE, from Vision to Road Map: Technical Steps

Mark Olive [a], Hanene Rahmouni [a], Tony Solomonides [a], Vincent Breton [b], Yannick Legré [b], Ignacio Blanquer [c] and Vicente Hernandez [c]

[a] *CCCS / CEMS Faculty / UWE, Bristol / Coldharbour Lane / Bristol BS16 1QY / UK*
[b] *LPC / CNRS-IN2P3 / Campus de Cézeaux / 63177 Aubière Cedex / France*
[c] *Universidad Politécnica de Valencia / Camino de Vera s/n / E-46022 Valencia / Spain*

Abstract.

We present the 'HealthGrid' initiative and briefly review work carried out in various European healthgrid projects. We report on joint work with numerous European collaborators. Since the European Commission's Information Society Technologies programme funded the first grid-based health and medical projects, the HealthGrid movement has flourished in Europe. Many projects have now been completed and 'HealthGrid' consulted a number of experts to compile and publish a 'White Paper' which establishes the foundations, potential scope and prospects of an approach to health informatics based on a grid infrastructure. With a second generation of projects now funded, the EC has commissioned the SHARE Project, a study to define a research roadmap for a 'healthgrid for Europe' as the preferred infrastructure for medical and health care projects in the European Research Area. The project explores the ways in which the healthgrid approach supports modern trends both in research in biomedicine and in healthcare, such as evidence-based practice and information integration.

Keywords:

healthgrid, e-health, grid applications

The HealthGrid initiative

'Grid' has been identified as one of the key technologies to support the European Research Area. The impact of this concept is expected to reach far beyond eScience, to eBusiness, eGovernment, and eHealth, but a major challenge is to take the technology out of the laboratory to the citizen. A *healthgrid* is an environment in which medical data can be stored and made available to all actors in the healthcare system, doctors, allied professions, healthcare centres, administrators and, of course, patients and citizens in general. Such an environment has to offer all appropriate guarantees in terms of data protection, respect for ethics and observance of regulations; it has to support the notion of 'duty of care' and may have to deal with 'freedom of information' issues. Working across member states, it may have to support negotiation and policy bridging.

Pioneering projects in the application of grid technologies to the health area have been completed, and the technology to address high level requirements in a grid environment has been under development and making good progress. Because these projects had a finite lifetime and the vision required a sustained effort over an extended period, and besides because there was an obvious need for these projects to cross-fertilise, the 'HealthGrid initiative', represented by the HealthGrid association (http://www.healthgrid.org), was launched to bring the necessary long-term continuity. Its goal is to encourage and support collaboration between autonomous projects in such a way as to ensure that requirements really are met and that the wheel, so to speak, is not re-invented repeatedly at the expense of other necessary work.

Writing about the healthgrid initiative very soon after its inception, this community identified a number of objectives [1]: identification of potential business models for medical grid applications; feedback to the grid development community on the requirements of the pilot applications deployed by the European projects; development of a systematic picture of the broad and specific requirements of physicians and other health workers when interacting with grid applications; dialogue with clinicians and those involved in medical research and grid development to determine potential pilots; interaction with clinicians and researchers to gain feedback from the pilots; interaction with all relevant parties concerning legal and ethical issues identified by the pilots; dissemination to the wider biomedical community on the outcome of the pilots; interaction and exchange of results with similar groups worldwide; and the formulation and specification of potential new applications with the help of the end user communities.

The grid concept is rooted in the physical sciences and these considerations were not a central concern to general grid developers. Even today these requirements are not a priority for developers, even though they have been fed through to the middleware services community. Thus HealthGrid identified the need for a specialist middleware layer, between the generic grid infrastructure and the medical or health applications.

Selected for best paper award.

Among data related requirements, the need for suitable access to biological and medical image data arose in several early projects, but for the most part these are present in other fields of application also. Looking to security requirements, most of these are special to the medical field: anonymous or private login to public and private databases; guaranteed privacy, including anonymisation, pseudonymisation and encryption as necessary; legal requirements, especially in relation to data protection, and dynamic negotiation of security and trust policies while applications remain live. Medical applications also require access to small data subsets, like image slices and model geometry. At the (batch) job level, medical applications need an understanding of job failure and means to retrieve the situation.

The white paper: from grid to HealthGrid

The next step for the HealthGrid community was to try to systematise the concepts, requirements, scope and possibilities of grid technology in the life sciences. The White Paper [2] defines the concept of a healthgrid more precisely than before: ... grid infrastructures comprising applications, services or middleware components that deal with the specific problems arising in the processing of biomedical data.

The ultimate goal for eHealth in Europe may be the creation of a single healthgrid incorporating a 'principle of subsidiarity' for independent nodes of the healthgrid as a means of implementing all the legal, ethical, regulatory and negotiation requirements. We may anticipate, however, the development path to proceed through specific healthgrids with perhaps rudimentary inter-grid interaction/interoperational capabilities. We may therefore identify a need to map future research and advice on research policy, so as to bring diverse initiatives to the point of convergence.

Healthgrid applications address both individualised healthcare – diagnosis and treatment - and epidemiology with a view to public health. Individualised healthcare is improved by the efficient and secure combination of immediate availability of personal clinical information and widespread availability of advanced services for diagnosis and therapy. Epidemiology healthgrids combine the information from a wide population to extract knowledge that can lead to the discovery of new correlations between symptoms, diseases, genetic features and other clinical data. With this broad range of application in mind, the issues below are identified as key features of our analysis.

- Business case, trust and continuity issues: healthgrids are data- and collaboration grids, but healthcare organisations are required by law to maintain control of their patients' records. Deployment on a scale to make an attractive business opportunity requires a high level security and compliance.
- Biomedical issues: Distributed databases and data mining are important tools for many biomedical applications in fields such epidemiology, drug design and even diagnosis. Expert system services running on the grid must be able to interrogate large distributed databases to explore sources of diseases, risk populations, evolution of diseases or suitable proteins to fight against specific diseases.
- Security issues: These flow naturally from the nature of medical data and from business requirements. Security in current grids is adequate only for research platforms.
- Management issues: The central concept of a 'virtual organisation' (VO) at the heart of eScience, which gave rise to grids, is very apt for healthgrid, but additional flexibility is needed to structure and to control VOs on a broader scale, including, for example, the meta-level of a VO of VOs.

We illustrate the concept of healthgrid with some prototypical examples: GEMSS [3] used a 'high-throughput' numerical simulation of organs obtained from a patients' data and used these to aid understanding or to improve the design of medical devices, with patient-customised approaches at research-level in areas such as radiotherapy, craniofacial surgery and neurosurgery. MammoGrid [4] created a database of standardised mammogram files and associated patient data to enable radiologists in the UK and in Italy to request second opinion and computer-aided detection services. The project also enabled a further study of an epidemiological nature on breast density as a risk factor. In Health-e-Child [5] radiologists are working with oncologists, cardiologists and rheumatologists to identify early imaging signs of conditions that may have a strong genetic component, possibly reducing the need for genetic maps to be obtained on an indiscriminate basis. In Wide In Silico Docking On Malaria (WISDOM) [6] tens of millions of molecular docking experiments have been used to help identify potential antigens for the malaria parasite. The experiment uses large scale virtual screening techniques to select molecular fragments for further investigation in the development of pharmaceuticals for neglected diseases. The economic dynamics in this area are telling: only about 1% of drugs developed in the last quarter century have been aimed at tropical diseases, and yet these are major killers in the third world, with mortality in excess of 14 million per annum. Meanwhile, the results of several major studies of the interface between bioinformatics and medical informatics had been published with a remarkable promise of synergy between the two disciplines, leading to what had already begun to be referred to as 'personalised medicine'. [7,8]

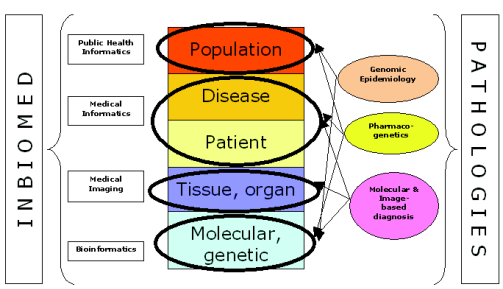


Figure 1 - Disciplines, levels of being and pathology diagnostics (F. Martin-Sánchez)

The SHARE Project: towards a road map

The vision of health that informs the thinking of the White Paper is reflected in European thinking [9] and is depicted in a map of the relationships between the different ontological and epistemological levels and the various modalities of data have been captured by Fernando Martin-Sánchez (cf [1]) in the schematic diagram of Figure 1.

In the White Paper, the HealthGrid community expressed its commitment to engage with and support modern trends in medical practice, especially 'evidence-based medicine' as an integrative principle, to be applied across the dimensions of individual through to public health, diagnosis through treatment to prevention, from molecules through cells, tissues and organs to individuals and populations.

In view of the impact of the White Paper, the EC has funded the project SHARE [10] to explore exactly what it would mean to realise the vision of the White Paper, investigate the issues that arise and define a roadmap for research and technology which would lead to wide deployment and adoption of healthgrids in the next ten years. Thus the project must address the questions, What research and development needs to be done now? and What are the right initiatives in eHealth RTD policy relating to grid deployment?, with all that implies in terms of coordination of strategy, programme funding and support for innovation. Thus the project will define a comprehensive European research and development roadmap, covering both policy and technology, to guide and promote beneficial EU-wide uptake of healthgrid technologies.

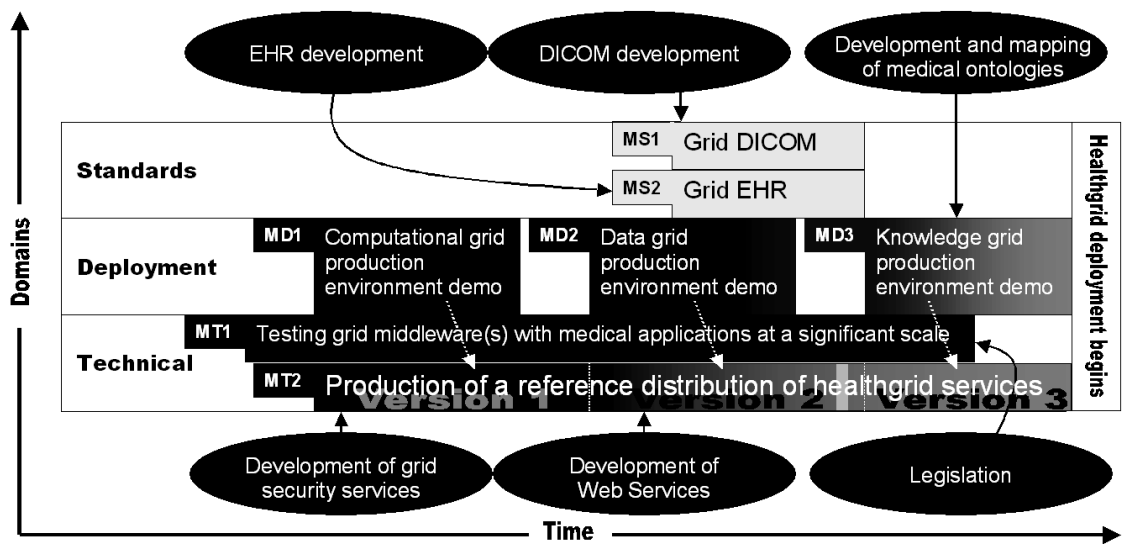


Figure 2 - SHARE technical roadmap diagram, showing milestones and external influences

Technical road map: Step one

SHARE has defined a preliminary technical road map (see Figure 2 above) with two technical milestones for appropriate development of healthgrid services (**MT1,2**), two milestones as examples of grid standards for the medical domain (**MS1,2**), and three deployment milestones of increasing complexity and scope (**MD1,2,3**).

MT1 Before deployment can begin, grid middleware must be tested with medical applications for scalability and robustness. This must begin at an early stage. It is anticipated that this will be an ongoing activity, with different generations of grid operating systems offering newer, faster and more stable capabilities.

A key issue for this milestone will be the robustness of grid solutions based on web services. Scalability, particularly regarding medical applications, is still a concern for grid middleware based on the Open Grid Services Architecture (OGSA) such as GT4 and GRIA. Middleware such as gLite and Unicore on the other hand have been deployed on large scale infrastructures in Europe and have demonstrated their scalability and robustness, but are still awaiting migration to web services.

MD1 The first deployment step will be the rollout of a computational grid production environment demonstrator for medical research. This would seem to be an achievable goal in the reasonably near future, given that there have already been successful deployments of computational grid applications (the WISDOM data challenges, for example) on general purpose grid infrastructures such as DEISA and EGEE. However, convincing healthcare management of the benefits of deploying a computational grid on a hospital or clinic IT infrastructure, which would not be composed of dedicated grid nodes and may already be working near capacity, is a real concern. Many medical centres may simply not have the necessary bandwidth or storage capabilities to make best use of grid technology, or may not have appropriate equipment to capture data in digital form. The management and configuration of a grid is rather complex and may require significant investment in manpower and training.

Installation of grid nodes behind hospital firewalls is incompatible with the present security model, which

required inbound connections. Secure services for data management are under development that could allow firewall rules to be relaxed, enabling connections to grid infrastructures. New architectures and designs should be defined that will minimise the volume of data leaving hospital borders.

Ideally, the grid node would be located outside the firewall with only anonymised or pseudonymised data being stored on the grid. However, even with the most stringent pseudonymisation and de-identification techniques there is still some risk of unauthorised re-identification by a person with sufficient knowledge from other sources. There are therefore legal and ethical implications of storing even anonymised personal data on the grid, and further investigation will be required to determine if this is possible with current national and European policies and legislation.

MT2 starts with **MD1** and ends when production deployment begins. This milestone is the development of a reference distribution of grid services, using standard web service technology and allowing secure manipulation of distributed data. Standards in emerging Web services technologies (WSx and the WSRF specification) will facilitate interoperability between healthgrids built using different underlying tools. A precise, well documented set of requirements is needed to describe the security features and obligation policies at different levels of abstraction in the middleware.

The IBHIS project found that web service description languages and registries are not yet mature enough, particularly WSDL, which describes how to access web services, and UDDI, which provides a registry for service discovery. These technologies are currently not flexible enough, cannot be used for semantic queries or descriptions (e.g. a description of the function a service provides, or the meaning of parameter names) or non-functional descriptions such as quality of service and performance levels. The development of WSDL and UDDI is ongoing, with OWL-S extensions to UDDI to facilitate semantic searching, upcoming versions of UDDI promising to address other limitations, and WSDL 2.0 promising to support semantic descriptions and include non-functional requirements.

Security is not an option but a mandate for healthgrids at all technical levels: networks need to provide protocols for secure data transfer, the grid infrastructure needs to provide secure mechanisms for access, authentication, and authorisation, as well as sites for secure data storage. The grid operating system needs to provide access control to individual files stored on the grid, and high level services need to properly manage the legal issues relating to the protection of medical data. Important progress is being made in terms of fine-grained access control and data encryption.

Revocation of credentials and how to provide temporary access to data is still an open issue, and an important one for healthgrids. There are a number of situations where users would temporarily require access to data that they would not normally have access to, such as a visiting expert being shown an unusual case. Certificate authorisation servers have been developed in both 'Pull' mode, in which sites periodically pull a list of valid members from a central service, and 'Push' mode, in which users obtain a short-lived attribute certificate that they present to sites to prove their membership. However, both of these would leave a window where revoked or expired credentials could be used to gain unauthorised access. Several healthgrid projects have suggested that the data itself should have a 'lifetime' – users with temporary access should not be able to access the data (or a copy of the data) once their credentials have expired.

MD2 Although several prototype data grids for medical research have been demonstrated by healthgrid projects, developing and maintaining a production quality data grid will require a number of issues relating to the distributed storage of medical data to be resolved. In European grid infrastructures, the distributed storage of medical images has been hampered by the limited data management services available, and so the continuation of improvements in this area will be important for the adoption of grids by the medical community. High speed links between data providers and consumers will be a prerequisite, particularly given the high volume of data predicted.

Many legal and ethical issues will need to be resolved, such as the ownership of patient data, ethical control of information, the patient's right to access or be informed about data that concerns them, as well as local, national, and European legislation governing the use of patent data and IT.

Another important concern for this milestone will be the integration of heterogeneous data from multiple sources. While mechanisms for data integration have been demonstrated by previous projects, biomedical data can be exceptionally varied including images with associated metadata and free form text or hand written notes from patient records. There is also the issue of how to deal with missing, inaccurate or obsolete data.

MS1 & **MS2** The use of computer-based tools for clinical research has led to the definition of standards for the exchange of data in many areas but their adoption has not been universal. The exchange of data between bioinformatics and medical informatics is an area where standards are particularly limited. By contrast, in medical imaging the adoption of DICOM for the storage and transmission of medical images has been accepted worldwide. In medical records HL7 is the emerging standard. For both of these standards, there is a question of compatibility with grid technologies.

MD3 After the issues with the distributed storage and querying of medical data have been resolved, the next task will be to deploy services that can build relationships between data items, and will provide appropriate representation to medical researchers. Particularly given that there have been no successful deployments of knowledge grids for medical research to date, this will pose a significant challenge. The data concerned can be extremely varied in nature, structure, format and volume. Depending on the

area of research, the synthesis of knowledge from data could require sophisticated data mining, integrated disease modelling and medical image processing applications, and may also involve the use of techniques from artificial intelligence to derive relationships between data from different sources and in different contexts.

The development of medical ontologies and mapping between ontologies will be particularly important for the successful deployment of knowledge grids. An ontology is the systematic description of a given phenomenon: it often includes a controlled vocabulary and relationships, captures nuances in meaning and enables knowledge sharing and reuse. From an agreed ontology it is possible to define a common data model that describes the format of the data used by all the services. The standardisation of interfaces can dramatically increase interoperability between biomedical resources, and by operating on standardised data formats they can more easily be integrated into complete bioinformatics experiments by eliminating the restructuring of data between each service. The construction of standardised data formats can be improved by defining a domain ontology that covers the concepts used within a given domain. These ontologies will allow relationships between concepts and nuances in meaning to be captured, greatly enhancing the opportunities for communication, knowledge sharing and reuse, and machine reasoning.

Open issues include how to integrate biomedical data using ontologies, how to combine different initiatives and how to employ advanced, semantic reasoning techniques for analysing medical data.

Conclusions and future work

Certain specific features of the community, such as issues of patient ownership of her/his data and the tension between hospitals' IT policies and the requirements of grids, will continue to prove troublesome unless addressed with political will. Another non-functional obstacle is the drag on technology transfer between EC projects. E.g. there is a need for healthgrid projects to begin thinking about data curation and digital libraries, but researchers and providers have not come together to explore this need.

SHARE predicts that it may take ten to fifteen years from a sustainable computing grid to a generalised knowledge grid. However, the transition to data grid may not be as simple as the success of special projects suggests and the transition to knowledge grids will be breaking new ground. It has been suggested that a more realistic timeframe might be twenty to forty years. As a next step, SHARE will focus on the large number of Ethical, Legal and Socio-Economic issues related to healthgrids. These will be integrated with the technical roadmap to recommend both technical and policy actions.

The work reported here has been carried out in collaboration with many colleagues in the HealthGrid community. Thanks are due to these and numerous other colleagues in Europe, the US and Asia for helpful discussions.

References

[1] V. Breton, A. E. Solomonides & R.H. McClatchey. *A Perspective on the Healthgrid Initiative*. Second International Workshop on Biomedical Computations on the Grid, the Fourth IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid 2004), Chicago, USA, April 2004

[2] V. Breton, K. Dean & T. Solomonides (Eds.) *The Healthgrid White Paper*. From Grid to Healthgrid: Proceedings of Healthgrid 2005: pp. 249-321, IOS Press, Studies in Health Technology and Informatics, Vol 112, 2005.

[3] GEMSS Project. *Grid Enabled Medical Simulation Services*. Available at: http://www.gemss.de

[4] C. del Frate et al. *Final Results and Exploitation Plans for MammoGrid*. Challenges and Opportunities of Healthgrids, Proceedings of HealthGrid 2006, IOS Press, Studies in Health Technology and Informatics Vol 120, 2006.

[5] J. Freund et al. *Health-e-Child: An Integrated Biomedical Platform for Grid-Based Paediatric Applications*. Challenges and Opportunities of Healthgrids, Proceedings of HealthGrid 2006, IOS Press, Studies in Health Technology and Informatics Vol 120, 2006.

[6] Jacq, N., Salzemann, J., Jacq, F., Legré, Y., Medernach, E., Montagnat, J., Maaß, A., Reichstadt, M., Schwichtenberg, H., Sridhar, M., Kasam, V., Zimmermann, M., Hofmann, M., Breton, V., *Grid-enabled Virtual Screening Against Malaria*, Journal of Grid Computing (to appear 2007).

[7] A. Sousa Pereira, V. Maojo, F. Martin-Sanchez, A. Babic & S. Goes. *The INFOGENMED Project*. ICBME 2002, Singapore, 2002.

[8] BIOINFOMED Project. *Synergy between Medical Informatics and Bioinformatics: Facilitating Genomic Medicine for Future Healthcare*, BIOINFOMED Study Report White Paper, 2003.

[9] European Commission. *e-Health – making healthcare for European citizens: An action plan for a European e-Health Area*. COM(2004) 356 final. A 'Communication from the Commission', available at: http://europa.eu.int/information_society/doc/qualif/health/COM_2004_0356_F_EN_ACTE.pdf

[10] V. Breton, I. Blanquer, V. Hernandez, Y. Legré & A. E. Solomonides. *Proposing a Road Map for Healthgrids*. Challenges and Opportunities of Healthgrids, Proceedings of HealthGrid 2006, IOS Press, Studies in Health Technology and Informatics Vol 120, 2006.

Address for correspondence

All correspondence to Tony Solomonides at Tony.Solomonides@uwe.ac.uk.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Penetration and Adoption of Health Information Technology (IT) in Thailand's Community Health Centers (CHCs): A National Survey

Boonchai Kijsanayotin[a], Supasit Pannarunothai[b], Stuart Speedie[a]

[a] *Health Informatics Graduate Program, Medical School, University of Minnesota, Minneapolis, Minnesota, USA*
[b] *Faculty of Medicine, Naresuan University, Phitsanulok, Thailand*

Abstract

A universal healthcare coverage program has been implemented in Thailand since 2001 and the Thailand Ministry of Public Health (MOPH) is restructuring its health information systems to support the management of this reform. The MOPH believes that health information technology (IT) is fundamental to the development of an effective health information system, and that users' adoption of health IT is one of the most important factors to the success of health IT implementation projects. However, there is no national data available regarding the penetration and adoption of health IT in Thai community health centers (CHCs). This cross sectional survey was designed to study the penetration and adoption of health IT in the country's community health centers. A random sample of 1,607 regionally stratified CHC's from a total of 9,806 CHCs was selected. With an 82% response rate, the data showed that people who worked in CHCs were currently heavy users of health IT. They exhibited high IT acceptance and positive attitudes toward using health IT. CHCs' staff was less resistant to adopt health IT than previously anticipated. These results are similar in all of the country's geographic regions. Health IT is pervasive in CHCs across the country and penetrates all regions.

Keywords:

community health centers, information systems, computer systems, attitude of health personnel, diffusion of innovation, Thailand

Introduction

The Thailand universal healthcare coverage program (the so-called 30 Baht for all diseases scheme) was instituted in 2001 and continues to evolve[1-3]. The program covers 74% of the 64 million people in Thailand. Thailand's Ministry of Public Health (MOPH) is currently restructuring its national health information system to support that reform. The MOPH is not only responsible for the country's public health system, but also is the country's major healthcare provider. In 2001, it oversaw 868 hospitals (67.1% of hospitals nationwide) and all 9,738 community health centers[4]. A community health center (CHC) is a subdistrict (Tambon) or village-level health service unit. This is a first-line unit, covering a population of about 1,000 - 5,000, with health staff including a health worker, a midwife and a technical nurse. Similar to other developing countries, the Thai government and MOPH administrators have recognized the potential of information technology to improve the quality of healthcare delivery and have embraced the technology[5].

Studies show that approximately 40% of information technology developments in various sectors, including the health sector, have been considered failures or have been abandoned. Moreover, this number has remained approximately the same for the last 25 years[6]. Littlejohns et al (2003) reported that the reasons for the failure of a large computerized health information system project in South Africa were a lack of users' understanding of reasons for new system and the underestimation of the complexity of the healthcare system[7]. Lorenzi and Riley (2003) pointed out that human issues at both individual and organizational levels are the reasons contributing to the information system (IS) implementation failures. They categorized reasons for information system failure as ineffective communication, underestimation of complexity, scope creep, organization problems, technology problems, and leadership issues. They emphasized the importance of organization change management when innovative information systems are being developed and implemented[8]. DeLone and McLean (2002) argued that information system usage is one of the six interdependent dimensions used to measure IS success[9]. The Heidelberg Health Information System (HIS) Working Group stated in the conference of the International Medical Informatics Association (IMIA) in 2003 that ".... people, not technology, will ultimately determine the success of HIS"[10]. Overall, these studies have concluded that the socio-technical aspect of the IT, particularly people and organizations, is essential to the success of the system[11].

With the advent of the universal healthcare coverage program, effective new health center information systems are anticipated in Thailand. The systems will link administrative, healthcare operations and public health information systems together in community health centers, local hospitals, and local and central administrative offices[12, 13]. Knowledge about users and organizational changes is one of the vital components for success of the system design, development and implementation[14,15]. However, national data regarding the current status of IT use in healthcare are not available. We do not know how healthcare providers currently use IT, or their willingness to accept IT in the CHCs. This knowledge is crucial for the

success of health information system development and implementation. Health IT in this study is defined as the information and communication technology that is used for health information systems. The present study had two main objectives. First we aims to describe the penetration of health IT in the country's CHCs and describe how CHCs currently use health IT, their attitude towards health IT and the degree of their acceptance of information technology in their work. Second, we aimed to inform health policy makers and those who are working in health IT projects about the readiness of community health centers to adopt and use health IT.

Methods

A cross sectional national survey was conducted in Thailand during July-October 2005. All 75 provinces except Bangkok, the capital, were stratified by four geographic regions; northern, central, northeastern and southern regions Then we randomly sampled population weighted provinces from each region. This resulted in the selection of twelve provinces with 1,607 CHCs: three from the north with 337 CHCs, four from the central with 376 CHCs, three from the northeast with 708 CHCs, and two from the south with 146 CHCs. At each CHC, we asked an officer who was responsible for the CHC's information management or the CHC's administrative officer (the CHC head officer) to complete the survey. Research collaborators at the provincial health office in each province distributed and collected the self-administered paper-based survey.

The survey instrument was developed by a group of Thai health IT experts. The instrument development was guided by two previous studies in which instruments had been tested and have exhibited acceptable levels of reliability and validity. Computer use and demographic measures were from the first author's previous work[16]. The attitude towards IT and intention to use IT measures were adopted from Venkatesh et al (2003)[17]. They were minimally modified and translated into Thai language so as to be clearly understood by Thai health center personnel. IT use associated with activities in CHCs was measured with ten questions using a four point scale ranging from one (never use) and four (always use). These ten items included four for "use for providing care and routine reporting", three for "use for management and administration" and another three for "use for information searching and collaboration with colleagues". CHCs' attitudes toward health IT were measured with 15 items representing four constructs: 1) performance expectancy (*PE*), 2) effort expectancy (*EE*), 3) social influence (*SI*) and 4) intention to use IT (*IN*), drawn from the Unified Theory of Acceptance and Use of Technology (UTAUT) model[17]. *PE* was defined as the degree to which an individual believes that using the system will help him or her to attain gains in job performance. *EE* was defined as the ease with which the system could be used. *SI* was defined as the degree to which an individual perceives that important others believe he or she should use the system. Intention to use IT was defined as the intention to use a computer system. *PE*, *EE* and *SI* each was measured with four items whereas *IN* was measured with three items. Each item was rated by a seven point scale ranging from one (the most negative perception) to seven (the most positive perception).

The pilot test was conducted with 36 part-time students of Faculty of Public Health at Naresuan University. These students are CHC officers from various parts of Thailand. SPSS V14.0 statistics package was used to perform descriptive statistics and ANOVA analysis.

Results

The response rate was 82% with 1,323 out of 1,607 CHCs responding. Of the 1,323 respondents, male and female were represented in essentially equal proportions (54% and 46%) and average age was 36 years old (SD=7, range 17-59). Three quarters (74%) of the respondents had a bachelor degree and one-third (32%) held a public health administration officer position. Nearly all respondents (99%) were in the middle level of government officer's classification position with 39% in the lower middle and 60% in the upper middle level. This means that they were not new employees but rather had worked in centers for several years, and had worked their way up through the position classification system. They spent aproximately 40% of their work hours providing health services and almost equal proportion of work hours in data management and report production. [Table 1] There was a slight difference in how the respondents were spending their work hours. CHCs in the south exhibited a slightly higher percentage of work hours in data management and reporting activities than the other regions.

Regarding IT resources, on average there were 3.7 persons (median = 3, range = 1-17) working in CHCs and most of these working persons (80%) used computer. [Table 1] Virtually all CHCs have at least one computer system. Only two CHCs (0.02%) reported that they did not have any computer system. On average there were approximately two computer workstations per health center and 36% of the CHCs had a local area network (LAN). Less than half of the CHCs could connect to the Internet and 83% of the connections were through a dial-up modem. Internet connectivity was slightly different in different parts of the country, with the highest proportion in the north and the lowest in the northeast. However, in the central region, the Internet connectivity of CHCs in the province adjacent to the capital, Bangkok was 93.36%. Despite only 47% of CHCs having Internet connectivity, 92% of the respondents reported that they had experience using the Internet. They may have used the Internet at a local Internet café, in educational institutions or at home. Moreover, there was no difference in Internet experience between regions. It is evident that IT is pervasive in CHCs and had diffused to all parts of the country.

Mean scores of the three different types of IT use constructs were 3.6 for "use for providing care and routine reporting", 3.2 for "use for management and administration" and 2.6 for "use for information searching and collaboration with colleagues." Cronbach alpha reliabilities were 0.75, 0.66 and 0.77 respectively. [Table 2] CHCs' personnel frequently used IT for providing care and producing reports, but used IT less in communication and information searching. Moreover, this pattern was similar across regions. This finding is consistent with the result

Table 1 - IT resources in CHCs and work hours spent by personnel in CHCs

Variables	Total	Central	North	Northeast	South	Statistics
1.Number of Community Health Center(CHCs)	1,607	376	377	708	146	
Number of response (% response)	1323(82%)	350(93%)	338(90%)	541(76%)	94(64%)	
2. Number of personnel /CHC - Mean	3.68	3.57	3.86	3.65	3.57	ANOVA *p=0.04* (df=3,1313)
3. Number of personnel using computer -Mean (% of number of personnel)	2.95 (80%)	2.98 (83%)	2.99 (77%)	2.9 (80%)	3.0 (84%)	ANOVA *p=0.54* (df=3,1220)
4. Computer systems in CHC (% of CHCs)	1321 (99.8%)	350 (100%)	338 (100%)	540 (99.8%)	93 (98.9%)	2, *p= 0.10 (df=3)*
Number of computer machines (PC, laptop)	2.10	1.91	2.11	2.16	2.19	ANOVA *p<0.01** (df=3,1305)
LAN (% of CHCs)	478 (36.1%)	113 (32.3%)	160 (47.3%)	175 (32.3%)	30 (31.9%)	2, *p<0.01* (df=3)
5. Internet connectivity in CHCs (% of CHCs)	621 (47.0%)	190 (54.3%)	215 (63.6%)	169 (31.3%)	47 (50%)	2, *p<0.01* (df=3)
6. Reported experience using the Internet	1214 (91.8%)	318 (90.9%)	318 (94.1%)	492 (90.9%)	86 (91.5%)	2, *p=0.35* (df=3)
7. Work hours spent (% of total work hours)						
Providing healthcare	39.1%	40.3%	37.2%	39.5%	39.2%	ANOVA *p=0.17* (df=3,1295)
Data management and reporting activities	39.8%	38.6%	41.4%	38.9%	44.0%	ANOVA *p<0.01** (df=3,1295)
Management & other activities	20.1%	21.0%	21.4%	21.6%	16.8%	ANOVA *p=0.04** (df=3,1295)

* Tukey HDS post hoc test also significant

that respondents spent 80% of their work hours providing health care services and performing data gathering and reporting activities. The use associated with information search and communication was relatively high in the north, which had higher Internet connectivity than other regions. The frequent use of health IT is confirmed by the finding that 92% of respondents reported that they have more than 3 years of experience using computers and two thirds of them use a computer more than one time per day. The results indicate that the computer is an integral part of the CHCs' work.

The mean score of *PE, EE, SI* and *IN* were 5.6, 5.2, 5.1 and 5.6 with the Cronbach alpha reliability 0.96, 0.95, 0.91 and 0.98 respectively. [Table 3] It appears that CHCs' personnel exhibited relatively strong positive attitudes towards health IT. They expressed high IT acceptance. In addition, the findings are similar among regions.

We also provided respondents with the opportunity to express their opinions and suggestions about health IT in CHCs. Half of the respondents (652 out of 1,323 respondents) provided comments. Content analysis of those comments reveals three themes. First, they perceived that IT is vital to their work environment. It helps them do their daily work. Second, they were overloaded with data management and reporting demanded from both local and central administrations. This result is supported by the finding that they spent almost half of their work hours doing the data and reporting activities. Sometimes they have to enter same data into multiple reporting systems. The final theme they perceived was that support from higher levels of the administration was inadequate. They suggested that local and central administrations should provide more technical support, IT training and expand infrastructure such as Internet connectivity and provide more frequent hardware upgrades.

Table 2 - IT use associated with activities in CHCs

Computer use	Total	Central	North	Northeast	South	Statistics
Use computer > 1 time /day : % of respondents	66.4% (866/1305)	55.6% (193/347)	73.1% (245/335)	70.1% (371/529)	60.6% (57/94)	2, *p<0.01* (df=3)
Have used computer ≥ 3 years : % of respondents	92.3% (1025/ 1306)	86.2% (299/347)	96.7% (326/337)	93.0% (491/528)	94.7% (89.94)	2, *p<0.01* (df=3)
Use for providing care and reporting activities: Mean score (SD)	3.6 (0.6)	3.5(0.5)	3.6(0.6)	3.5(0.6)	3.4(0.7)	ANOVA *p=0.06* (df 3,1198)
Use for management and administration activities: Mean score (SD)	3.2(0.7)	3.1(0.7)	3.2(0.6)	3.2(0.6)	3.0(0.6)	ANOVA *p=0.01* (df=3,1055)
Use for information searching and collaboration with colleagues (e.g. email) : Mean score (SD)	2.6(0.8)	2.5(0.7)	2.9(0.7)	2.6(0.8)	2.5(0.8)	ANOVA *p<0.01** (df=3,651)

Measurement scale 1 = Never use, 2 = Sometimes use, 3 = Frequent use and 4 = Always use
* Tukey HDS post hoc test also significant

Table 3 - Attitudes towards Health IT

Attitudes	Total	Central	North	Northeast	South	Statistics
Performance Expectancy (*PE*) Mean (SD)	5.6 (1.8)	5.7 (1.7)	5.6 (1.8)	5.6 (2.0)	5.4 (2.0)	ANOVA *p=0.56 (df=3,1284)*
Effort Expectancy (*EE*) Mean (SD)	5.2 (1.6)	5.2 (1.6)	5.2 (1.6)	5.2 (1.7)	5.0 (1.6)	ANOVA *p=0.62 (df=3,1279)*
Social Influence (*SI*) Mean (SD)	5.1 (1.7)	5.2 (1.6)	5.1 (1.6)	5.1 (1.8)	5.0 (1.5)	ANOVA *p=0.73 (df=3,1283)*
Intention to use IT (*IN*) Mean (SD)	5.6 (1.8)	5.8 (1.7)	5.6 (1.8)	5.6 (1.9)	5.5 (1.9)	ANOVA *p=0.44 (df=3,1292)*

Measurement scale 1 = Strongly Disagree, 2 = Quite Disagree, 3 = Slight Disagree, 4 = Neither Agree nor Disagree, 5 = Slight Agree, 6 = Quite Agree and 7 = Strongly Agree

Discussion

It appears that computer and communication technology is currently in heavy use by CHCs across the country. Information technology utilization penetrates and diffuses to CHCs in every region. CHCs have adopted health IT as part of their work. These results contrast to the previous perception of Ministry of Public Health that CHCs might be slow to adopt health IT because majority of CHCs are in rural areas. People who work in CHCs applied computer technology to conduct and enhance the health services they provide and reporting tasks they perform. This might result from the influence both from within MOPH and from rapid social change toward globalization in Thai society as a whole. Since the universal coverage scheme was implemented, both local and central health administrators have increasingly demanded numerous timely reports of health activities for planning and management. These demands come from reporting systems from various departments and institutions with minimal integration or interaction – typical silos. Utilizing a computer system enables CHCs to accommodate such demands with less effort than they might otherwise need. Relatively high positive attitudes toward health IT among CHCs indicates their high degree of IT acceptance. They demonstrate high receptivity to the use of IT in healthcare. The fact that the majority of the respondents were young with college degrees, and have been exposed to computer technology for years might be a significant contributing factor to the high IT acceptance. However, the study findings illustrate other challenges that might threaten the success of health IT project implementation. The success of the systems

might be threatened by connectivity problems, if the country's goal is to develop integrated health information systems, since over half of the CHCs do not have any ability to connect directly to the Internet. Bandwidth is also a potential problem, since most of those CHCs that are able to connect are using dial-up as the mode of access. The communication infrastructure to all CHCs needs to be expanded. The other challenges are related to the data management and reporting overload including the perception of inadequate IT training and IT resources from the higher level of the organization. These issues should be considered for further studies.

Although this survey was designed to be representative of all CHCs nationwide, our assumption that the respondents represented the CHC (which is the unit of analysis) might be a potential weakness. However, given the consistency of the responses we do not believe the results have been compromised by this assumption or its occasional violation.

Conclusions

Thai health administrators and policy makers anticipate having an effective information system to support the healthcare system similar to administrations in others countries. Lessons learned from previous national health technology implementation projects suggest that user acceptance is the major determinant of project success. We surveyed a representative sample of Thai CHCs nationwide to find out the extent of IT penetration among CHCs and to evaluate their attitude towards and acceptance to IT. The data demonstrate that health IT is pervasive and there is a high degree of IT acceptance and use in CHCs across the country. The study results suggest that many of the necessary socio-technical conditions are already in place, and that the potential for their interference with a successful implementation of the national health information system is smaller than anticipated. However, the inadequate Internet connectivity and the data management and reporting overloads are the challenges that might interfere with the success of health IT projects.

Acknowledgements

The authors gratefully acknowledge Dr.Suvat Siasiriwattana, Mrs Aree Wadwongtham, Dr.Narong Kasitipradith and the staff at the Bureau of Health Policy and Strategy, Ministry of Public Health, Thailand for their great support.

The authors thank Ms.Parunee Yimsabai and staff at the Center of Health Equity Monitoring (CHEM), Faculty of Medicine, Naresuan University, Thailand for helping the survey process. The primary funding source of this project was from WHO Thailand, project activities proposal number 040199.

References

[1] Pannarunothai S, Patmasiriwat D, and Srithamrongsawat S, Universal health coverage in Thailand: ideas for reform and policy struggling, Health Policy, 2004. 68(1): p. 17-30.

[2] Tangcharoensathien V, Pitayarangsarit S, and Vasavid C, Universal health care coverage and medium term financing implications., International Health Policy Program-Thailand: Nonthaburi, 2002

[3] Towse A, Mills A, and Tangcharoensathien V, Learning from Thailand's health reforms. BMJ, 2004. 328(7431): p. 103-105.

[4] Wibulpolprasert S, Siasiriwattana S, Ekachampaka P, Wattanamano S, Taverat R (Editor), Thailand Health Profile 2001-2004, Bureau of Policy and Strategy, Ministry of Public Health.Thailand,2005

[5] Chandrasekhar CP and Ghosh J, Information and communication technologies and health in low income countries: the potential and the constraints, Bull World Health Organ, 2001. 79(9): p. 850-5.

[6] Kaplan B, Shaw NT, Future directions in evaluation research: people, organizational, and social issues, Methods of Information in Medicine, 2004. 43(3): p. 215-31.

[7] Littlejohns P, Wyatt JC, and Garvican L, Evaluating computerised health information systems: hard lessons still to be learnt, BMJ, 2003. 326(7394): p. 860-3.

[8] Lorenzi NM and Riley RT, Organizational issues = change, Inter J Med Inf, 2003. 69(2-3): p. 197-203.

[9] DeLone WH and McLean ER, Information systems success revisited, Proceeding of the 35th Annual Hawaii International Conference on System Sciences (CD ROM) 2002.

[10] Giuse DA and Kuhn KA, Health information systems challenges: the Heidelberg conference and the future, Inter J Med Inf, 2003. 69(2-3): p. 105-114.

[11] Whetton S, Health informatics: a socio-technical perspective, Oxford; New York: Oxford University Press. xxv, 262 p. 2005

[12] Information and Communication Technology Center, Ministry of Public Health, Ministry of Public Health's Information and Communication Technology Master Plan 2004-2006(Thai Language). [Web page] [cited November 29, 2006]; Available from: www.morph.go.th/ict/data/MOPH_ICTPlan.zip.

[13] Bureau of Policy and Strategy, Ministry of Public Health, Health Policy in Thailand 2006, Ministry of Public Health Thailand. [Web page] [cited November 29, 2006]; Available from: http://203.157.19.191/HealthPolicy6.pdf.

[14] Lorenzi NM, Riley RT, Blyth AJC, Southon G, Dixon BJ, Antecedents of the people and organizational of medical informatics: Review of the literature, J Am Med Inform Assoc, 1997. 4(2): p. 79-93.

[15] Dixon DR, *The behavioral side of information technology.* Int J Med Inf, 1999. 56(1-3): p. 117-23.

[16] Kijsanayotin B., Pannarunothai S, and Speedie S, Development of an instrument to measure health enter (HC) personnel's computer use, knowledge and functionality demand for HC computerized information systems in Thailand, Proceeding of the AMIA 2005 Annual Symposium(CD ROM), Washington DC, 2005

[17] Venkatesh V, Morris MG, Davis GB and Davis FD, User acceptance of information technology: Toward a unified view. MIS Quarterly, 2003. 27(3): p. 425-478.

Address for correspondence

Boonchai Kijsanayotin MD., MS.
Health Informatics Graduate Program,
Medical School,
University of Minnesota Twin Cities,
420 Delaware, Minneapolis, MN 55455
Email: kijs0001@umn.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

The Health Informatics Center of Acadiana – Informing Health Policymaking in Post-Katrina/Rita Louisiana

L. Philip Caillouet Ph.D.[a]

[a]*Health Informatics Center of Acadiana, The University of Louisiana at Lafayette, Lafayette, Louisiana, USA*

Abstract

A "healthy communities" initiative in Louisiana led to creation of the Health Informatics Center of Acadiana (HICA) at The University of Louisiana at Lafayette, in the south central United States. Since hurricanes Katrina and Rita devastated the Louisiana coast in 2005, HICA's role has taken on heightened significance. HICA identifies vulnerable populations, documents their risk factors, and evaluates interventions intended to improve community health. HICA collaborates with the Louisiana Department of Health and Hospitals and the Lafayette Community Health Consortium (LCHC), the latter formed for coordination among local healthcare providers and agencies. Both HICA and LCHC were created when "Bonne Santé à Lafayette!" – a locally developed community health improvement plan – was implemented. This paper reports on methods and experiences of HICA and LCHC, offering these as models for addressing community concerns elsewhere. Of special interest is the discussion of Louisiana HABITS, a consumer survey methodology that HICA has developed to measure healthcare access barriers, to provide information that healthcare organizations and governments need to implement workable business strategy and public policy.

Keywords:

public health informatics, health policy, access to health care, health informatics, public health

Problem: Poor health or poor health policy?

Health insurance and healthcare in the USA

In the pluralistic United States healthcare system, healthcare services are offered principally by private-sector professionals and institutions rather than through government agents. These entities are sustained with cash flow margins generated by market-driven compensation for their services. A wide variety of health insurance arrangements reduce the financial risks of prospective patients with varying benefit definitions, limits of coverage, and modes of oversight. Most U. S. residents obtain health insurance coverage for themselves and their dependents through employer-sponsored plans, the premium costs of which are most often shared between the employee and the employer. This somewhat haphazard scheme has been entrenched since the period immediately following World War II, but in recent years a steady increase in the number of uninsured Americans has been of concern to policy makers and ordinary citizens alike. Many federal agencies are involved in the enforcement of public health policy and management of categorical problems related to disease control and prevention, health resource distribution, and health promotion – the latter currently through the U. S. Surgeon General's initiative called Healthy People 2010. [1] However, local management of those programs is largely left to the individual states, as is the dilemma of what to do about the uninsured and under-insured.

Health insurance and healthcare in Louisiana

In the year 2000, Louisiana had a population of 4.5 million. [2] According to data published for that year [3], approximately 19.1% of Louisiana citizens lacked health insurance in 2000. These persons might not have been able to find a private provider willing to treat them unless they are able to pay out-of-pocket for their healthcare costs. An additional 16.2% of the population was insured through the Medicaid program in the same year, meaning that their income and resources were meager. Under Medicaid, state governments administer and share costs with the federal government for health insurance coverage for the qualifying poor. Because Medicaid payments to providers are not always competitive with those from other forms of coverage, providers are often reluctant to see Medicaid patients. Compounding the problem, it is increasingly difficult to recruit and retain private providers willing to locate in both rural and urban areas where pockets of poverty or near poverty exist. Thus the uninsured and under-insured in Louisiana, some 35.3% of its residents, are under-served.

While some federal funding is channeled into qualified community health centers (FQHCs), Rural Health Clinics (RHCs), and Critical Access Hospitals (CAHs) in Louisiana, because of the large number of under-served persons Louisiana has felt it necessary to complement services of private providers by offering both ambulatory and in-patient care directly. Louisiana's governments have had a long history of intervention on behalf of the health of its residents. In the eighteenth and nineteenth centuries, Louisiana's sub-tropical climate was an especially good host to a variety of diseases including yellow fever and malaria, which during periodic outbreaks reached epidemic levels. On March 15, 1855, Louisiana became the first state to enact legislation creating a permanent State Board of

Health. Currently, the Louisiana Department of Health and Hospitals (DHH) is responsible for both the Medicaid program and for the direct care network. Public health services are managed centrally at the state level, with nine regional administrators being responsible for implementation of programs. Region size varies from four to twelve parishes (a "civil parish" is the Louisiana equivalent of a county), with regional populations ranging from 280,000 to 1.2 million. [2] Relative autonomy in public health services administration is exercised locally in the city of New Orleans but not in any other locale. Organized within DHH, the Office of Public Health (OPH) operates at least one clinic or "health unit" in each of Louisiana's sixty-four parishes. These health units provide limited disease prevention and health promotion services as mandated through categorically funded governmental programs. Louisiana is the only one of the fifty United States to have a state-owned, state-operated hospital system. The "Charity System" takes its name from one of the oldest continuously operating hospitals in the United States. Charity Hospital in New Orleans – now part of the Medical Center of Louisiana – was founded in French colonial Louisiana in 1736, sixty-seven years before the United States negotiated the Louisiana Purchase. Until the late summer 2005 flooding caused by Hurricane Katrina rendered its facilities unusable, "Big Charity" was the flagship institution of a geographically dispersed including ten hospitals managed through the Louisiana State University Health Care Services Division.

Despite this long history the state appears to be mired in an extended cycle of poor outcomes, with at least one observer ranking Louisiana as having one of the United States' worst records of relative healthiness. [4] It is widely debated as to whether Louisiana's approach represents a progressive mechanism for dealing with the problem of the insured or is merely a remnant of a failed, two-tiered, paternalistic "plantation system," as some privately describe it. In recent years, a new climate of urgency has arisen in Louisiana that, with the full support of the State Health Officer, seeks to encourage communities and regions within Louisiana to define local needs and seek local solutions – without necessarily relying on state government to take the initiative. Because Louisiana has under-invested in health assessment, a consistent theme in these local initiatives has been the lack of actionable information on which to base effective policy changes.

An informed-leadership approach

Community initiatives in Lafayette, Louisiana

Lafayette – the "capital of Acadiana" – is the hub of south Louisiana's famous Cajun culture, which highly values self-sufficiency and independence, but at the same time embodies an enlightened view of community responsibility among physicians and other healthcare providers. Yet, as the commercial center of the fourth largest metropolitan statistical area in Louisiana, Lafayette shares in Louisiana's poor showing among states. A signal event did occur however in 1994 when two competing not-for-profit community hospitals in Lafayette, Lafayette General Medical Center and Our Lady of Lourdes Regional Medical Center, jointly sponsored and conducted a Community Needs Assessment. Out of that project, the Partnership for a Healthier Lafayette ("the Partnership") was later chartered as a "healthy communities" planning group in 1996. The Partnership was formed as a joint project of the United Way of Acadiana, the Greater Lafayette Chamber of Commerce, the Lafayette Economic Development Authority, and Lafayette Consolidated (City-Parish) Government. OPH Region IV in south central Louisiana was a key player in from the inception of the Partnership. The Partnership's chosen methodology was to engage community stakeholders in envisioning the future and identifying what it would take to realize that vision. "Trend-bender Teams" were formed to propose solutions not only in Healthcare but also in the areas of Education, Youth & Family, Teen Pregnancy, Economic Development, Civic Participation & Volunteerism, Community Infrastructure, and Transportation. The report of the Healthcare Trend-bender Team was completed in December, 1997, and published as part of *A Vision for Our Future* in April, 1998. [5]

Besides the Partnership for a Healthier Lafayette, several other healthy community programs were being launched in Louisiana about the same time. Most notably, the Turning Point Partnership project, sponsored under the auspices of OPH, attracted funding from the Robert Wood Johnson Foundation and the W. K. Kellogg Foundation to support the development of Public Health Improvement Plans. Strategic planning has proceeded at the statewide level and in with the participation of local organizations known as "New Orleans, the City That Cares", "the Southwest Louisiana Partnership" (Region V), and "the Northeast Louisiana Partnership" (Region VIII). [6] Unlike in Turning Point, the planning approach used in Lafayette has focused not so much on improving "public health" (i.e., the services of the state agency) and has instead focused on improving "the health of the public." During the local planning sessions, this subtle distinction evolved into a gradual recognition that the traditional methods of data collection and reporting by public health agencies in Louisiana are inappropriate to reporting on community health status, let alone to enabling decision making for measurable enhancements in community health status. It is to the credit of local public health professionals that they were among the first to agree with this notion when it first surfaced.

Informed leadership via health informatics

In Lafayette, a particular emphasis on "informed leadership" emerged. The rationale for the Health Informatics Center of Acadiana ("the Center" or "HICA") stems directly from the development of *Bonne Santé à Lafayette!* (Good Health to Lafayette!), the community health status enhancement initiative of the PHL Healthcare Trend-bender Team. The *Bonne Santé à Lafayette!* program incorporates both structural and action-oriented elements intended to focus the attention of the healthcare providers on identification and solution of community health problems. The two major critically interdependent structural elements of *Bonne Santé à Lafayette!* are HICA and the

Lafayette Community Health Consortium ("the Consortium" or "LCHC"). While the latter mobilizes healthcare community leadership, the former provides a means of informing the decisions and actions of those leaders. To be truly effective however, the LCHC must exercise *informed* leadership, not just leadership. That is where the Center enters this picture.

Creation of the Health Informatics Center of Acadiana

HICA was established at the University of Louisiana at Lafayette in February 1999, in conjunction with the Department of Health Information Management and the Health Care Administration MBA program. The Center acts as a local clearinghouse for to be used in documenting community health status, in identifying vulnerable populations, and in studying risk factors within these groups.

Mission of the Health Informatics center of Acadiana

The mission of HICA is four-fold:

1. Education of health professionals and healthcare administrators at undergraduate, graduate, and continuing education levels;
2. Research into community health needs and into effectiveness of the healthcare community's response;
3. Health status enhancement in Louisiana, by serving as a resource to healthy communities initiatives; and
4. Health policy enhancement in Louisiana, by serving as a resource to policymakers related to healthcare access, delivery, and financing.

In pursuit of this mission, HICA leverages both long-standing health information management principles and emerging information and communications technologies to collect, aggregate, analyze, and report community health information. The common need for actionable information cries out for the acceptance of consensus standards in the area of computer applications, communications infrastructure, data definitions, and for sharing of best operational practices. HICA serves as a focal point for that development, and for the implementation of appropriate community health information networks, one goal being the release of community health status report cards.

Activities of the Health Informatics Center of Acadiana

HICA gathers, analyzes, and disseminates information essential to identifying and prioritizing community health needs. HICA was created to address many of the concepts of a "management information system for community health" and will ultimately rely on information networking technologies to achieve its mission. Local empowerment and the facilitation of community priority setting should be greatly enhanced in comparison to the current approach of shipping data on locally addressable issues to state and federal agencies only to have it return as out-of-date printed statistics! HICA has even being viewed as "good for business," [7] in terms of its potential for fostering an environment where the community can mobilize to eradicate inequalities in access to healthcare and thereby a healthier labor force. A strong relationship is evolving between HICA and emerging "healthy communities" programs in the area, particularly with the Bayou Teche Community Health Network (ByNet) and the Vermilion Community Health Network (VNET). HICA is working closely in academic-industry partnerships with other groups, including the Emergency Medical Services (EMS) Council of the Lafayette Area, to find funding for projects to obtain, aggregate, analyze, and report on health services utilization data in Region IV.

Related academic programs

The Center also complements the healthcare-related educational and research missions of the University while functioning as a resource to the healthcare community in the immediate area and throughout the state. In Lafayette, both The University of Louisiana at Lafayette (UL Lafayette) and Louisiana State University Health Sciences Center (LSUHSC) each prepare students in the health professions, allied-health fields, and health care administration locally. UL Lafayette currently provides baccalaureate curricula in Nursing and Health Information Management (BS), plus master's degrees Business Administration (MBA) with a Health Care Administration option and Nursing (MSN), with additional preparation for as nurse practitioner (NP) certification. LSUHSC conducts postdoctoral (MD) residency programs for recent graduates. Members of the UL Lafayette faculty have also participated in the continuing education of public health professionals in various settings. [8] HICA serves as a focal point for interdisciplinary curricular and research activities for students and faculty and expects to be an integral aspect of a proposed interdisciplinary graduate program in Health Informatics.

Results – research themes & projects

Research at the Health Informatics Center

Research at HICA has taken several directions. Community health enhancement issues are most clearly associated with what can be called public health informatics. Since Louisiana ranks at or near the bottom of entirely too many categories within which public health is commonly measured, many research questions come to mind immediately. What scales were used? What measurements were taken? Can Louisiana really be that bad off? What are other states doing to better their lot? How could Louisiana get off the bottom of the list? Graduate student research papers have already begun to examine these questions. Louisiana is very fertile ground for advanced research in the public health informatics field, and the existence of HICA will aid in reaping a significant harvest, appealing especially to faculty and students in health care administration, communications, marketing, biostatistics, epidemiology, and public policy.

Louisiana HABITS

HICA is gradually becoming a clearinghouse for data to be used in documenting community health status, identifying vulnerable populations, and studying risk factors in need of remediation within these groups. Reliable parish-by-parish health risk data has not been generally available in

Louisiana. While OPH does publish parish statistics based on results of the Centers for Disease Control (CDC)-sponsored national Behavioral Risk Factor Surveillance System (BRFSS), but the small sample size (limited by data collection budgets) of fewer than 140 random-dialed statewide interviews per month does not lend meaningful statistical significance to findings at the parish level. Given the unique blend of ethnicities and cultures in Acadiana, the applicability of these findings to this area was called into question.

Since March 1999, a HICA research team has worked on the refinement and deployment of HICA's own methodology for assessing healthcare access barriers, real or perceived by healthcare consumers. The tool and approach were dubbed *Louisiana HABITS* (Healthcare Access Barriers In The State). *Louisiana HABITS* has been applied in support of seven distinct health assessment projects supported directly or indirectly by the Rapides Foundation, the Robert Wood Johnson Foundation, and the LSU Foundation/Pfizer Inc.

In its planning to understand consumer perceptions and demand for healthcare services, the *Louisiana HABITS* team defined healthcare access barrier to mean difficulty, delay, or failure in obtaining healthcare or prescribed medications in the past twelve months, or current lack of health insurance, experienced by one or more family members in a household. The team took on the task of answering the question: How should the percentage of all households in the general population that have healthcare access barriers be most easily and accurately determined? Fundamental concerns included statistical validity and protection of respondent confidentiality. The team recognized that random-digit-dialed telephone interviews would be preferred, due to low cost when compared to in-person interviewing and high compliance when compared to mailed surveys.

The *Louisiana HABITS* team reviewed the most current U. S. Census Bureau data, to determine the population and number of households and to set criteria for random sample size sufficient to yield 95% predictive confidence, with a maximum error rate of ±10%. A very costly fifteen-fold increase in interviews would have been necessary to have achieved a ±2.5% interval. The UL Lafayette-developed computer-assisted *Louisiana HABITS* consumer survey was then employed to gather data from a random sample of at least 96 households with telephones, to determine the proportion of the general population of households that report having a healthcare access barrier.

The *Louisiana HABITS* team was concerned, however, that prior surveys using a telephone-only interviews were inaccurate when a certain fraction of all households have no telephone. A review of 1990 U. S. Census data suggested that a substantial fraction of households (e.g., 11.3% of households in St. Mary Parish, Louisiana) were without working telephones at that time. Although Census 2000 data was not available at the time of the first survey, the team felt it reasonable to assume that the proportion of households without telephones had likely fallen due to the increased use of cellular telephones in the decade of the 1990s.

To find a representative sampling of no-phone households, the *Louisiana HABITS* team conducted in-person interviews at locations in each parish where persons whose households might not have telephones could be readily found. In-person interviewing locations included social services offices and parish health units, and occasionally included other public places such as courthouses, public hospital emergency rooms and clinics, and even rural grocery stores and laundromats. As expected, a strong positive correlation was quickly noted between households without telephones and households with healthcare access barriers, with financial reasons as the cause.

In-person interviewing was continued in each parish until data were obtained from at least 96 households in the barrier population in each parish. Demographic stratifiers such as age, gender, marital status, and education level, and general health of the healthcare decision maker were also obtained, as well as household income, household size and composition, and insurance coverage status. Thus in addition to allowing the computation of a no phone adjustment to the findings of the random-digit-dialed telephone survey, the *Louisiana HABITS* team also gained, through in-person interviewing, statistically significant predictive knowledge of the underlying causes of the barriers, including the following:

- Main reason cited by those reporting a problem obtaining healthcare services,
- Main reason cited by those reporting a problem obtaining prescribed medications, and
- Main reason cited by those reporting lack of insurance, and other pertinent statistics.

Knowledge of these underlying causes for healthcare access barriers has already been helpful in the design of interventions in each the twenty-two parishes where *Louisiana HABITS* has been conducted – Acadia, Allen, Avoyelles, Beauregard, Calcasieu, Cameron, Catahoula, Concordia, Evangeline, Grant, Iberia, Jefferson Davis, Lafayette, La Salle, Natchitoches, Rapides, St. Landry, St. Martin, St. Mary, Vermilion, Vernon, and Winn.

Other notable activities

During the several difficult months of population displacement that followed hurricanes Katrina and Rita in the fall of 2005, HICA promoted deployment of an electronic health record for use with evacuees in shelters, but encountered obstacles due to absence of specific authority within the emergency response structure. Since then HICA has assisted in the reshaping of policy at the local, state, and regional levels with a view to making health information technology more responsive to the needs of the victims of such disasters.

HICA has also undertaken partnerships with public and private health-related agencies to act as data collector and analyst on significant federally funding initiatives:

- Process and outcome evaluation for "Healthy Start in Lafayette Parish," working with The Family Tree – funded by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services.
- Process and outcome evaluation for "Safe Schools / Healthy Students," working with the Lafayette Parish School System – funded jointly by the U.S. Departments of Education, Justice, and Health and Human Services.
- Process and outcome evaluation for the "Lafayette Jail Diversion Program," working with the Louisiana Office of Mental Health – funded by the Substance Abuse and Mental Health Services Administration (SAMHSA) of the U.S. Department of Health and Human Services.
- Data collection for the Behavioral Risk Factor Surveillance System (BRFSS), in partnership with the Louisiana Office of Public Health – funded by the Centers for Disease Control and Prevention (CDC) of the U.S. Department of Health and Human Services.

At this writing, HICA is in discussions with the Louisiana Department of Health and Hospitals to act as architect for an Aging and Adult Services single-point-of-entry system. Two additional data collection contracts are also pending between HICA and the Louisiana Public Health Institute (LPHI) for the following projects:

- Data collection for the 2007 Louisiana Adult Tobacco Survey, funded by the Centers for Disease Control and Prevention (CDC) of the U.S. Department of Health and Human Services.
- Data collection for the 2007 Steps for a Healthier New Orleans BRFSS, funded by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services.

Conclusions

The synergy catalyzed by the presentation of credible community health information in readily usable form to committed, decisive community healthcare leadership is expected energize future efforts to enhancing its health status for years to come! Much remains to be learned as the Health Informatics Center of Acadiana evolves. Perhaps this model of collaboration among academic, community, industry, and government resources will be replicated in other settings where similar concerns exist.

References

[1] U. S. Department of Health and Human Services. Healthy People 2010: Understanding and Improving Health. Washington: U. S. Government Printing Office; 2000.

[2] U. S. Census Bureau. Ranking Tables for Counties: Population in 2000 and Population, Change from 1990 to 2000 (PHC-T-4). Available at: http://blue.census.gov/population/www/cen2000/phc-t4.html. Accessed July 3, 2001.

[3] Louisiana Department of Health and Hospitals, David W. Hood, Secretary. 2002 Louisiana Health Report Card, Submitted to the Governor and the Louisiana Legislature. April 29, 2002.

[4] ReliaStar Financial Corporation. The ReliaStar State Health Rankings – An Analysis of the Relative Healthiness of the Populations in All 50 States, 1997 Edition. 1997.

[5] Partnership for a Healthier Lafayette. A Vision for Our Future. April 1998.

[6] Turning Point -- Louisiana Style: Thinking Differently, Doing Things in a New Way. In: Transformations in Public Health, a publication of Turning Point: Collaborating for a New Century in Public Health. vol 1, issue 2, June, 1998, pp. 5-6, 8.

[7] Health Informatics Center Proposed. In: Lafayette Business Journal, May, 1998. p. 13.

[8] Partnership for a Healthier Lafayette: A Model for Louisiana's Healthy Communities Initiatives, Presented at: Louisiana Public Health Association 1998 Educational Conference, Lafayette LA, April 16, 1998.

Address for correspondence

L. Philip Caillouet Ph.D., Associate Professor & Director
Health Informatics Center of Acadiana
Post Office Box 41007
The University of Louisiana at Lafayette
Lafayette LA USA 70504-1007
E-Mail: caillouet@louisiana.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Informatics Solutions for Emergency Planning and Response

Elizabeth E. Weiner [a], Patricia A. Trangenstein [a]

[a] *Frist Nursing Informatics Center, School of Nursing, Vanderbilt University, Nashville, TN, USA*

Abstract

Early informatics contributions to the emergency planning and response agenda have focused largely on surveillance of threat detection. A broader assessment of possible informatics contributions unveils that informatics can also contribute to increasing the efficiency in disaster response as well as providing a tele-presence for remote medical caregivers. This presentation will explore current and future roles of informatics in emergency preparedness and response.

Special challenges for data management occur with every emergency or disaster. Tracking of victims, electronic health records, and supply inventory are a few of the contributions that informatics can play during disasters. Modeling of response resources can provide the parameters for more effective decision making. Public relations reporting can be made more accurate if given the information in a timely fashion. Databases provide the infrastructure for reporting of data that can be used to manage volunteers or later be mined to determine the effectiveness of planning and response efforts. As informaticists, we have a moral obligation to contribute to the emergency response agenda worldwide.

Keywords:

disaster response, emergency response, emergency planning

Introduction

All disasters begin at a local level, but some situations exacerbate into regional, national, or international events. The complexities of disaster event management are difficult enough, but are magnified when international response is deemed necessary. According to the United Nations International Strategy for Disaster Reduction (UN/ISDR), 2005 saw an 18% rise in natural disasters.[1] An estimated 160 million people – seven million more than in 2004 – were directly affected by natural disasters.

Natural disasters are not the only threats to world health. Currently, there are around 20 significant ongoing armed conflicts throughout the world. [2] Events of terrorism are reported on a daily basis internationally. In addition, there are manmade and technological disasters. The result is that our planning and response efforts need to be improved at all levels, including better preparation and resource management.

One aspect of resource management is that of human resources. Effective planning and response teams need to be reflective of a broad skill set of professionals who can best contribute. One important contribution can be that of the informaticist. Teich, Wagner, Mackenzie, and Schafer noted that with disaster, terrorism, and response to war, "the need for applied informatics expertise may be more pressing, and more in the public eye, than ever before." [3]

The purpose of this paper is to present current informatics contributions and explore possible future solutions to the challenges brought about by emergency planning and response.

Issues and discussion

Biosurveillance and bio-agent detection

The threat of bioterrorism has brought greater attention to the public health infrastructure in all countries. Public health leaders have long complained that there was an inadequate infrastructure to comply with simple public health demands. When that system was further threatened with a biological attack, the breakdowns in the infrastructure became readily apparent. The most noticeable weakness was the relationship between hospitals and public health organizations. Many times there was minimal reporting of possible trends or disease patterns found in hospitals that might possibly have an impact on all of public health. Healthcare workers are now aware of the implications and consequences of their clinical decisions with respect to the entire community. Informatics has played a major role in designing ongoing systematic collection, analysis, and dissemination of data about disease. While some hospitals continue to use manual tracking systems, the best solutions are those that occur in real time and thus immediately identify areas of concern.

Besides the need for additional funding for public health infrastructure, an additional need has been for policy changes in how information is handled. One such example is the Unified Medical Language System from the US National Library of Medicine. This standardized vocabulary tool helps to share descriptions across vocabularies and even link a new bioterrorism-monitoring vocabulary to other terminologies. [4] The synergy between standardized clinical data models and electronic health records have allowed clinicians to use the Internet to rapidly implement large-scale, multi-institutional clinical data integration.

Lober, Karras, Wagner, et al., led a roundtable discussion during AMIA 2001 to discuss and compare six existing bioterrorism detection systems.[5] Although the systems were developed independently, they had striking similarities in the types of data collected and the overall system architectures. All sites indicated concerns with maintaining security and confidentiality. Most used encryption for data transmission. Several systems used clustering of data codes to define disease prodromes of interest in bioterrorism detection. Differences in the systems related to the fact that the projects represented differing relationships with the public health system. As a result, some systems collected multiple levels of data, while others used the visit or case-report level of detail. Continued work in biosurveillance systems has continued since 2001 with funding in the US from the Centers for Disease Control and Prevention and the Agency for Healthcare Research and Quality.

Increasing efficiency in disaster response

Information sharing

Informatics solutions provide the possibility for improving the speed and quality of information that is shared between and among organizations responding to emergency and disaster events. Of particular importance is connecting those in the field who are directly responding to events with those organizations such as hospitals that will be receiving event victims. One example is Maryland's communication network (known as the Trauma Line), which enables prehospital field care providers to communicate directly with physicians in trauma centers and other referral centers. Data on patient vital signs, estimated time of arrival and means of transport, type of injury, level of consciousness, and priority status is put on a fax notepad linked to a cell phone in the ambulance for transmission to the hospital trauma team [3] Another project, know as "MobiDoc," makes use of wireless technology to create an entirely mobile telecommunication system.[3] A field team can perform multiple charting, vital-signs monitoring, image collection, and other data acquisition tasks for multiple patients. The data are sent securely to the hospital's intranet. More patient data and arrival information coupled with algorithms to make use of the data to balance resources will allow for more efficient care, including a need to increase capacity in local areas.

Even the reporting of victims during a mass casualty event has created challenges. After the attacks of September 11, thousands of family members circulated throughout the hospitals in the area in a futile attempt to locate their family members. There was not one central place for them to access the information. Healthcare members in St. Louis wanted to make certain this did not happen to their community. As a result, they developed a bar code system to log and track their victims. [6] In addition, PDAs were used by medics to log patients and belongings as well as notebook computers with wireless technology and networked desktop machines in command centers.

A report of the 2002 Coastal North Carolina Domestic Preparedness Training Exercise described the innovative use of telehealth technologies for terrorism response. [7] During this exercise, East Carolina University tested the in-place telehealth networks as well as deployable communications, networking, and data collection technologies such as satellite communication, local wireless networking, on-scene video, and clinical and environmental data acquisition and telemetry. Specific recommendations were shared based on their experience.

Information sharing during the response to Hurricane Katrina was hampered as it related to supply needs and inventory. For example, a call would go out that bottled water was needed in a certain region. There was no centralized method to determine if those needs were met and by whom, or how long it was going to take to deliver the needed goods. Inventory codes varied by vendors, so that comparing resources across vendor product lines was essentially impossible.

Information sharing during a disaster is essential as it relates to the media. The media provide the interface to the public at large, and it is important that the information they receive is both accurate, timely, and from the designated authority. Secure communication methods need to be established prior to an event. Appropriate informatics tools can help to make that happen.

Preparation (competency-based learning and simulations/exercises)

Provider training and education are also critical elements of a comprehensive plan for bioterrorism and public health preparedness in general. Researchers at Vanderbilt University have developed online modules for nurses in emergency planning and response. In addition to being competency-based from the International Nursing Coalition for Mass Casualty Education, they also used the "How People Learn" (HPL) format. This format was based on a review of the literature for the National Research Council on how people best learn. Current modules can be found at the following address free of charge: www.incmce.org

Other US funding sources have provided for additional learning materials. The CDC supports the linking of a number of academic institutions to state and local health agencies. The result is the Centers of Public Health Preparedness (CPHP) program (http://www.bt.cdc.gov/training/cphp/). The American Medical Association supports its National Disaster Life Support (NDLS) Program, which is also competency based across levels or expertise (http://www.ama-assn.org/ama/pub/category/6206.html)

In Japan, centers of excellence are also funded and recognized for their expertise in emergency planning and response. Hyogo University is such a designated center, focusing on disaster nursing in a ubiquitous society. Central to their efforts is a research focus addressing the development of an information base, the establishment of a nursing support network throughout Asia and the world, and the development of nursing care strategies which includes the Education/Training Program Development Project. [8]

A unique collaboration in emergency preparedness results in a Masters' degree. Called the European Master in Disas-

ter Medicine, The University of Eastern Piedmont School of Medicine, Novara, Italy, and the Free University of Brussels School of Medicine, Brussels, Belgium, together are the founding universities and actually provide the degree. [9] Associate universities include the Centre for Teaching & Research in Disaster Medicine and Traumatology (Linkoping, Sweden), Centre Hospitalier Universitaire Vaudois (Lausanne, Switzerland), the International Emergency Medicine and Disaster Medicine Sections, Harvard Medical School (Boston, Massachusetts, United States of American [USA]), Yale New Haven Center for Emergency Preparedness and Disaster Response (New Haven, Connecticut, USA), and the Vanderbilt University School of Nursing and National Center for Emergency Preparedness (Nashville, Tennessee, USA). Didactic content is presented in an online format for three months prior to the two week live-in course venue. The live-in venue takes place in a selected site in Italy, where interactive exercises culminate in a community-based drill.

Developers at Dartmouth's Interactive Media Laboratory have developed the Virtual Terrorism Response Academy (VTRA)(http://iml.dartmouth.edu/vtra/). The VTRA is a reusable virtual learning environment to prepare emergency responders to handle high-risk, low-frequency events, partiicularly terrorist attacks.[10] Users have the opportunity to sign in according to their profession and get a mentor/host from their designated profession. The most advanced simulations have been funded and used by the US military. The ultimate simulation environment will allow for individuals to participate in teams, preferably team members with whom they will respond to events. Informatics solutions will allow for feedback to be provided to both individuals and teams.

Volunteers

Healthcare volunteers are a necessary component of disaster and emergency response, but they also create challenges. Issues related to registration and credentialing can be solved with informatics participation. The design of the database is only the first challenge, followed by political issues such as who owns the data and how the data is collected, stored, and shared. Scopes of practice are governed by different groups regionally, nationally, and internationally. Liability issues surface with every response.

One example of volunteer registration and credentialing is the Emergency System for Advance Registration of Volunteer Health Professionals (ESAR-VHP). [11] This US based system is a state-based registry of volunteer health professionals. It allows for verification of credentials prior to a disaster, and provides opportunities for education and training in disaster response. In addition, volunteers are urged to register within one and only one registry so that duplicate records are not generated. It is hoped that this practice will discourage the spontaneous volunteer [12]. The success of any response is dependent on all individuals or teams understanding their roles, responsibilities, and the chain of command within the disaster response. Disaster respondents who deploy within a system that is part of the overall disaster planning usually come with the essential equipment that is needed to complete their mission and to keep the responder safe from hazards that might result from the disaster. A spontaneous volunteer who shows up on the disaster scene may not have the equipment, knowledge, skills, or team work skills to be able to be used in a successful response effort.

Electronic health records

Not all victims of disaster and emergency events have electronic records during non-emergency times. However, for those who do, access to such records could streamline the time needed to administer healthcare. Questions also arise during emergency events as to who owns the record? Particularly for those victims that are displaced, should the record follow the person or go to the next healthcare provider? How will either healthcare provider be identified in order to access a centralized database?

Hurricane Katrina provided such a challenge. Dr. David Brailer, (who was then the National Coordinator for Health Information Technology) pulled together informatics specialists from both academic and vendor backgrounds. These experts created www.KatrinaHealth.org, an online service for authorized health professionals. [13] The web site provided access to evacuees' medication information in order to renew prescriptions, prescribe new medications, and coordinate care. This website provided authorized users with access to the medication history of evacuees who lived in the areas affected by Hurricane Katrina, with data or prescription information made available from a variety of government and commercial sources. Sources included electronic databases from community pharmacies, government health insurance programs such as Medicaid, private insurers, and pharmacy benefits managers in the states most affected by the storm. In less than three weeks after Hurricane Katrina hit the Gulf Coast, this information was up and running. Developers agreed that it should not have taken a major disaster to stimulate this sort of collaborative work. As a result of these efforts, KatrinaHealth.org won the 2006 Pinnacle Award from the American Pharmacists Association Foundation.

The Markle Foundation convened a group of industry and government experts following Hurricane Katrina who prepared the summary report which included the need to:

- Foster immediate discussions regionally and nationally among government health leaders, insurers, healthcare providers, and information technology companies to determine what, how, and when patient medical information can be shared securely and quickly in the event of a disaster.
- Create electronic health information systems that are based on simple, open web standards, so that data can be provided in different formats from different users and still be accessible to all.
- Agree upon a method to authenticate the identities of doctors, pharmacists, other health professionals, and patients using the web site, so that they can quickly and securely access private health information needed for their ongoing treatment.

- Make electronic health information records accessible to nurse practitioners, physician assistants, and nurses who will likely be working with physicians and clinics in a disaster's aftermath, rather than just by physicians.
- Examine federal and state public policies governing privacy and medical records-such as the Health Insurance Portability and Accountability Act of 1996 and existing state privacy laws-to be sure they do not hinder the delivery of medical care for displaced persons post-disaster. [13]

Modeling

The establishment of an Emergency Operations Center (EOC) is one of the first steps towards a positive disaster response. Depending on the nature of the disaster situation, a collection of individuals are gathered to respond and plan next steps of the response effort. Based on standardized job action sheets, these leaders each play an assigned role in order to provide a comprehensive based for effective decision making. Recognizing the importance of technology, some EOCs now include a Technology Coordinator at high levels of decision making.

Although the event begins at the local level, regional, national, and even international responders can be called in to respond. However, it may be as long as 48-72 hours before additional help can actually occur on-site. In the interim, members of the EOC need to be able to make decisions regarding priorities, resources, and next steps. Some factors are known prior to an event, such as how long it will take before pharmaceutical stock piles can be brought on-site. Models that include such "known" entities can be used by these decision makers so that more efficient decisions can be made, perhaps saving lives due to the decreased time. Our colleagues in nuclear engineering have taken advantage of such modeling techniques, which can be adapted for use in both public health and healthcare organizations.

Telehealth applications

Expert assistance may not always be available on-site during a disaster. One method to provide this expertise is through telemedicine technologies. This enables an expert in one location to direct field personnel on-site. It is also feasible for that expert to provide assistance to more than one site.

Remote management of trauma patients also allows for remote guidance of procedures. Teich et.al. found that the use of tele-mentors took longer than if the expert were on site, but concluded that the tele-mentor was better than having no trauma surgeon present at all. [3]

Telehealth applications rely on having an intact communications infrastructure. In some situations, the infrastructure is destroyed, leaving clinicians to resort to less technological solutions.

Possible informatics research

The discussion above has generated a number of areas for further informatics research as related to emergency planning and response. These include the following:

- Efficient development and delivery of more advanced biosurveillance systems
- Development of more accurate algorithms for evaluating epidemiological data
- Continued work on vocabulary standardization
- Design of more regional or national databases related to the electronic health record
- Effectiveness of available online learning resources
- Development of enhanced telecommunication/telehealth technologies
- Development of data mining techniques that would assist in determining the effectiveness of response efforts
- Development and evaluation of appropriate modeling techniques to enhance the efficiency of decisions made during response situations

Conclusions and summary

Effective application of informatics to the challenges brought about by disasters can greatly enhance planning and response efforts. Examples of the appropriate use of informatics have been provided in the areas of biosurveillance, efficiency of response, and telehealth. The responsibility for the inclusion of informatics lies with each individual and organization who can lend expertise. Only then can we continue to play a prominent role in the prevention and management of disaster and emergency events.

Acknowledgments

Thanks to the US Department of Health and Human Services for their continued funding and support for the International Nursing Coalition for Mass Casualty Education, which is hosted by the School of Nursing at Vanderbilt.

References

[1] United Nations Fact Sheet. 2006. Retrieved December 3, 2006 from http://www.unisdr.org/eng/public_aware/world_camp/2006-2007/pdf/Fact-sheet-en.pdf

[2] WHO. The first consultation on nursing and midwifery contributions to emergencies. 2006. Available at: www.who.int/hac/events/22_23November2006/en/index.html.

[3] Teich JM, Wagner MM, Mackenzie CF, Schafer KO. The informatics response in disaster, terrorism, and war . J Am Med Inform Assoc. 2002; 9(2):97-104.

[4] Lober WB, Karras BT, Wagner MM, et al. Roundtable on bioterrorism detection: information system-based surveilance. J Am Med Inform Assoc. 2002; 9(2): 105-115.

[5] Kohane IS. The contributions of biomedical informatics to the fight against bioterrorism. J Am Med Inform Assoc. 2002; 9(2):116-119.

[6] Hamilton J. An internet-based bar code tracking system: Coordination of confusion at mass casualty incidents. Disaster Management & Response 2003; 1(1):25-30.

[7] Simmons SC, Murphy TA, Blanarovicii A, Workman F, Rosenthal DA & Carbone M. Telehealth technologies and applications for terrorism response: A report of the 2002 Coastal North Carolina domestic preparedness training exercise. J Am Med Inform Assoc. 2003; 20: 166-176.

[8] Yamamoto A. Disaster nursing in a ubiquitous society. Japan Journal of Nursing Science 2004; 1(1): 57-63.
[9] European Master in Disaster Medicine. Retrieved December 3, 2006 from http://www.dismedmaster.com/
[10] Dartmouth Interactive Media Laboratory. Retrieved December 3, 2006 from http://iml.dartmouth.edu/vtra/
[11] Peterson C. Be safe, be prepared: Emergency system for advance registration of volunteer health professionals in disaster response. OJIN2006: 3(2): Retrieved December 3, 2006 from www.nursingworld.org/ojin/topic31/tpc31_2.htm
[12] American Hospital Association. Emergency system for advanced registration of volunteer health care personnel – Hospital implementation issues and solutions focus group meeting report. 2005. Retrieved December 2, 2006 from http://www.hrsa.gov/esarvhp/FocusGroupReport0705/default.htm
[13] KatrinaHealth. Website for organization 2005. Retrieved December 2, 2006 from http://www.katrinahealth.org/
News release. Website for KatrinaHealth.org. 2006. Retrieved December 2, 2006 from (http://www.markle.org/resources/press_center/press_releases/2006/press_release_06132006.php)

Address for correspondence

Betsy.Weiner@vanderbilt.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

A Multidiscipline Conceptual Framework for Consumer Health Informatics

Robert A. Logan[a], Tony Tse [ab]

[a] *Lister Hill National Center for Biomedical Communications, National Library of Medicine, NIH, Bethesda, MD, USA*
[b] *Clinical Research Policy Analysis and Coordination Program, NIH, Bethesda, MD, USA (current affiliation)*

Abstract

This paper presents an idealized conceptual framework for consumer health informatics research drawing from complementary disciplines: information science and health campaign research. This synthesis is designed to provide researchers with a flexible model to evaluate current research and inform future studies. Following a description of the major components, we describe a recent evaluation of consumer perceptions of a health information system, Genetics Health Reference. This study illustrates how the framework may be applied to provide some direction and insights into ongoing consumer health informatics research. While this model represents a work in progress, we present it in support of efforts to understand the multidimensional impacts of the public's access to health information. We also discuss challenges that remain to develop a better conceptual understanding of how consumers converge on health informatics services.

Keywords:

health behavior; internet; informatics; communication; information science; consumer satisfaction; evaluation studies

Introduction

Napoli [1] suggests that consumer health informatics research underemphasizes the importance of comprehensive theoretical underpinnings in evaluating how consumers interact with health information services. Napoli [1] encourages informatics researchers to:

- Account for an array of institutional, social, professional, individual challenges that impede interactive consumer-based health informatics.
- Propose comprehensive models that underlie and instruct how research might be conceptually conceived and ultimately conducted.
- Borrow theoretical frameworks from health campaign research in conjunction with more traditional informatics research, such as information sciences.

Napoli [1] and Dutta Bergman [2] agree that a challenge in health campaign and consumer health informatics research is to provide a more holistic conceptual foundation, which would help scholars conceive and critique both existing and future research efforts. Health campaigns are non-commercial, research-based efforts that seek to change a target audience's awareness or basic knowledge about a disease, condition, or public health issue [2].

Baker and Pettigrew [3] explain a conceptual framework also serves as a roadmap, which helps a researcher conceive projects in a multidimensional fashion. Reinforcing the idea that there is nothing more practical than good theory, a broad conceptual framework identifies the areas that should be considered in planning an evaluation and clarifies the conceptual omissions that need to be explained.

In this paper, we present an idealized conceptual framework that integrates components from information science and health campaign research. A recent evaluation of a consumer health information Website is used to illustrate how the application of the framework provides some direction and adds insights in ongoing consumer health informatics research.

Our objective is to "sketch" the conceptual landscape representing and integrating existing theories about consumer interaction with health information media. While this model represents a work in progress, we present it in support of efforts to understand the multidimensional impacts of the public's access to health information and account for the related infrastructural, biopsychosocial, and interactional dimensions that have been identified – especially in health campaign research.

Background

In this section, we summarize foundational concepts from two disciplines that have explored information seeking from two different perspectives: (1) information science and (2) health campaign research.

Information science

The well-established information science literature conceptualizes the information-seeking process (ISP) as a dynamic, iterative series of cognitive processes and physical actions required for satisfying information needs (e.g., [4-7]). Several "core" non-linear, dynamic, and iterative information-seeking states are common to these ISP models:

- Need: identifying/expressing information needs;
- Access: finding relevant information; and
- Evaluation: assessing information that is found.

Briefly, information needs represent gaps in knowledge. The process of seeking "missing" knowledge helps to resolve these needs. However, consumers may have trouble identifying missing knowledge (i.e., knowing what is not known). Challenges include ineffective query formulation resulting from a lack of knowledge about medical concepts and terminologies [8] as well as finding credible information resources, constructing well-formed mental representations, and navigating online systems. The retrieved information is evaluated within the context of the original problem, another potential barrier for consumers in the medical domain.

Thus, from a traditional information science perspective, ISP consists of making sense of information within a dynamic environment. While these models typically emphasize decision-making and other cognitive processes, very few conceptual frameworks accommodate well-identified variables within the health campaign literature, including affective dimensions of consumer behavior [6] and the socio-cultural and economic environment that surrounds media use [7,2].

Health campaign research

Although health campaign researchers have embraced an individual-centered model, in recent years they have been prone to discuss the phenomena of how persons interact with health media (e.g., health news and information sources on the Internet), as a process of convergence rather than information seeking [9]. The term *convergence* is used because it is perceived to bridge both consumer information seeking and an array of more affective-oriented rationales for why people use mass media, often collectively referred to as *gratifications* [10].

Overall, health campaign research theory recognizes an array of interactions that conceptually frame the dynamics that occur in consumer health informatics. These dynamic interactions encompass intrapersonal, interpersonal, demographic, and cultural factors, as well as media source credibility, preparation of messages, and channel characteristics [2].

McGuire [9] and Cappella [11] note that the conceptual development of health campaign research has occurred in three phases. The first phase is similar to the ISP model described previously and emphasizes optimizing an individual's exposure to resources with their information needs [9].

The second phase emphasizes the potential communication barriers inherent in the characteristics of how a health media or interpersonal source is perceived. It encompasses potential communication barriers presented by messages, media channel, receiver (individual) and destination characteristics [10]. In other words, health messages may not be optimally understood by consumers if the message is poorly written, if a written rather than a visual media is used to reach some audiences, if a person has little access to mass media, or if messages are poorly timed with an individual's or target audience's needs. The emphasis in the second phase is on both individual perception and media characteristics [9].

The third phase encompasses: a) immediate social influences, such as the influences of peer pressure and commercial advertising on health behaviors, b) cognitive behavioral factors, such as a person's problem and decision making skills, and c) the degree that a commitment to a specific health behavior requires broader skills and individualized training to foster a healthier lifestyle. This phase includes theoretical models as well as approaches to research health message effects [11]. These include behavior change theories, individual information processing modes, message effects research and immediate systemic factors. It also emphasizes a person's individual cognitions and skills as well as his or her behavioral adaptability and milieu, or immediate social surroundings [11].

Dutta Bergman [2] adds that health campaign research has entered a fourth phase where the emphasis is to introduce macro forces, such as a nation's or region's health resources, its global or national context (e.g. developing versus industrialized, cultural and religious heritage) its economic prosperity and geo-political factors. Macro influences are seen as relevant to a conceptual understanding why health campaigns are accepted or rejected by intended audiences.

Dutta Bergman [2] notes a challenge of the fourth phase is to integrate all four dimensions described above. The model proposed below is one approach for integrating the interaction of the phases of health campaign research with the ISP model.

A multidimensional framework

Drawing from a new generation of relevant health campaign research and other ISP models and theories, we propose a conceptual framework (Figure 1) for visualizing high-level health information dimensions that span and bridge two disciplines. Each major dimension is described briefly and encompasses the literature cited above. The model also includes an important added dimension that we term 'outcome.'

Consumer/individual

This dimension covers how a person responds to health communication messages and notes psychological, motivational, immediate social, family factors, and related applied context. This is discussed in the description of phases one, two, and three in health communication research above as well as ISP research. Psychosocial factors may influence the basic information-seeking process (described previously). For instance, an individual's cognitive abilities, affective state, life skills and existing knowledge of the information problem, domain, or information source is likely to interact with the perception of information need, motivation, or effort spent evaluating retrieved information.

Other variables include individual demographic attributes, social influences, personal goals, and how a person perceives a health message. For example, a message that seems incon-sistent with an individual's personal beliefs (i.e., cognitive dissonance) may contribute to ending an

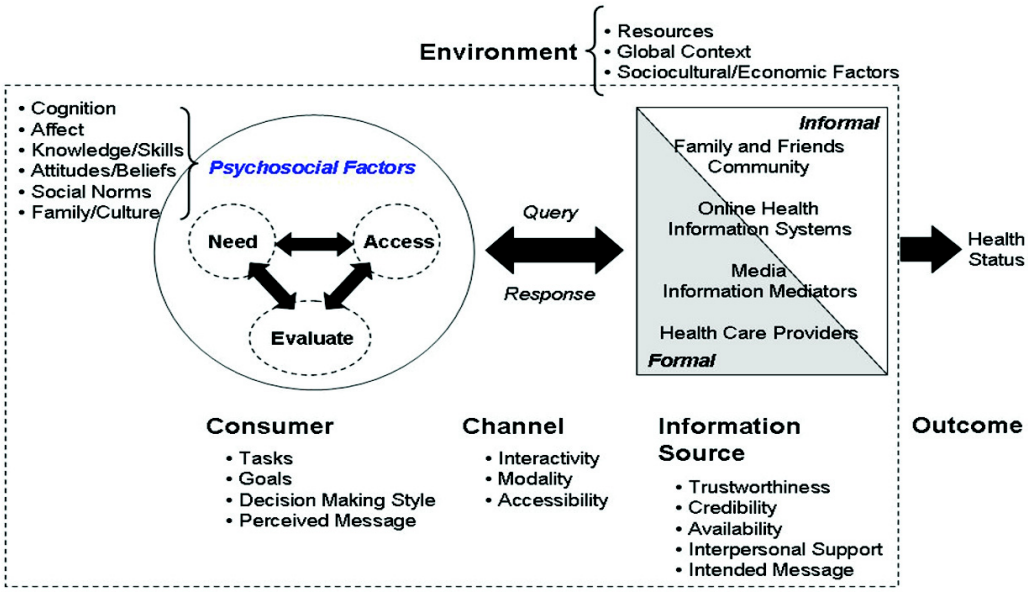


Figure 1 - Conceptual framework for consumer health informatics with five interactive dimensions – Consume (or Individual), Channel, Information Sources, (Macro) Environment, and Outcome

information-seeking session before an information need has been met.

Channel

This dimension describes the influence of the media channel by which health messages are conveyed. This is discussed in the description of phase two of health campaign research above. Channel includes attributes such as the level of interactivity (e.g., unidirectional versus bidirectional), modality (visual, audio, or multimodal), and accessibility (e.g., due to noise or insufficient bandwidth). For instance, some consumers prefer multimedia compared to text (only).

Information sources

This dimension encompasses the perception of the credibility of personal, professional, and media sources and how it influences an individual's reception of health messages. This is discussed in the description of phase two and three of health campaign research above. Many formal and informal sources of health information are available to consumers. Since information sources are typically characterized by perceptions of trustworthiness, credibility, availability, intrapersonal support, and the intended message, health communication from multiple formal (health care providers) versus informal sources (media, health informatics Websites, family and friends) can be a source of confusion for individuals.

Macro environment

This dimension includes the socio-economic, and cultural environment or heritage in which health communication occurs. This is discussed in phase four of health campaign research above. The context, such as available resources due to market and economic forces and socio-cultural trends and traditions, influences all of the other dimensions. For example, the "digital divide" and the impact of religious beliefs impact how individuals perceive health messages.

Outcome

Although health information convergence results in many outcomes (e.g., satisfying an information need), a critical, elusive component is finding therapeutic associations between consumer information access and resulting health behaviors:

> Despite abundant speculation regarding the consequences of consumer participation in interactive health communication, little research has investigated these issues… Ultimately interest and research on effects should focus on quality of health and health care. [12:686-7]

This dimension was added to the above description of ISP and health campaign research.

Overall, each component and its attributes potentially affect health information seeking and convergence on individual behaviors. People do not seek health information in a vacuum—it is integrated into their lives. How and when the need to seek information (perceived or real) overcomes other competing needs, interests, and activities may be based on predictable factors, such as risk-benefit analysis, or simple serendipity. Nevertheless, the design of effective consumer health information systems will likely improve significantly only after researchers better understand the practical role and nature of health information convergence on individual behaviors.

Sample application

While we recognize the complexity of the model and the research challenges posed by multiple interacting components and dimensions, we believe that the framework is useful for presenting a high-level map of areas to be explored. As Greenberg and colleagues state, "the more we know about [online consumer health information seeking] variables, the better we can design educational and technical strategies that help consumers get to the information they seek" [13:1].

Recently, the proposed framework was applied in evaluating the National Library of Medicine (NLM) site, Genetics Home Reference (GHR) [14]. GHR helps consumers understand information about genetic conditions and their related gene or chromosome variations.

A survey was conducted to evaluate perceptions of GHR from consumers' perspectives. Between February and April 2004, 374 members of the Genetic Alliance, an international advocacy group for people with genetic conditions, completed online surveys designed to probe multiple dimensions of consumer perception: content, design, and interface. The survey and data collection were approved by the U.S. Office of Management and Budget.

After consideration, two dimensions (macro environment, channel) were bypassed for investigation because of an anticipated high acceptability of health information seeking on the Internet among the study's participants. Two dimensions (individual and information sources) were pursued. Within the latter dimensions, researchers focused on investigating the perceived uses and gratifications of the Website by exploring participants' thoughts (cognitive) and feelings (affective) about GHR. Researchers also wished to assess participants' assessment of GHR's credibility. The focus on uses and gratifications resulted in questions about the Website's aesthetic appeal and emotions associated with using GHR [15]. Perceived affective dimensions, source credibility, and cognitive uses were derived from semantic differential scales, which are often used in studies of audience media perceptions [9].

By including a range of variables within the individual and information source dimensions, researchers were able to assess whether content quality and aesthetic characteristics in addition to traditional measures such as demographics, online experience, interest and ease of site navigation, predicted user satisfaction.

The study found that content quality and affective dimensions each predicted overall consumer satisfaction [15]. However, age, gender, prior online experience, interest, and education did not predict consumer satisfaction, which is contrary to some previous findings [16].

Further, in a factor analysis that combined all the 13 outcome variables that assessed cognitive, affective feelings and source credibility, the researchers found there were three distinctive perceptual orientations towards GHR [15]:

- Visual design and appeal
- Perceived source credibility/information quality
- Perceived complexity/simplicity and potential bias

In short, participant motivations to use GHR were more holistic than an expected interest to retrieve high quality information about genetics.

The point is that, by using the conceptual framework and integrating variables based on its dimensions, findings about the relative importance of how persons project attitudes onto GHR were better identified. The conceptual framework broadened the research variables that were pursued. In turn, these yielded results that expanded the existing literature about consumer motivations to use a consumer health Website. By initially focusing on a broader conceptual framework, the investigators defined narrow regions to explore and areas to bypass, as well as identified new types of research orientations that yielded surprising results.

Challenges

The health campaign research literature notes an array of methodological challenges, such as understanding audience segmentation and accounting for macroscopic sociopolitical and economic influences on the behavior of both consumers and media organizations [2]. Similarly in information science, there are challenges to capturing the "personal value" consumers place on health information, understanding relevance to particular health needs in the context of situational variables, and tracking the episodic and often serendipitous nature of health information seeking among formal and informal sources. Finally, a dearth of common terms across disciplines and validated instruments for measuring variables hinders consumer health informatics research [17].

Initial methodological challenges we encountered in applying a comprehensive framework include:

- Difficulty in obtaining representative samples online, including underserved populations
- Challenges in identifying and creating operational definitions, and isolating key variables such as "information exposure"
- Lack of standard assessment methods, variables, and operational definitions across research studies

Nevertheless, idealized, comprehensive conceptual frameworks are important in assisting investigators critique the dimensions they have encompassed or eclipsed in modeling consumer informatics research. While the model presented here reinforces this suggestion, it also underscores that even a consumer-centered, psychosocial approach represents only one of the major components of a more expansive, interactive system. It is important to evaluate both motivations for consumer behaviors and operant factors—information source, channel, consumer and environmental. The latter may be useful to explain why intended audiences are drawn or repelled to health Websites and embrace or reject health management and information-seeking concepts in the first place.

Conclusion

Although we may never fully account for the multidimensional spectrum that an idealized model represents, a comprehensive framework adds accountability to social research that fosters consideration of a range of issues and encourages investigator disclosure of the dimensions that are less explored.

An idealized conceptual framework outlines the considerations for which researchers should strive, debate, and defend. A model needs to be comprehensive in order to be ultimately useful and it needs to set a conceptual tradition that is well-grounded in sister disciplines. We hope such a model of consumer information seeking and convergence has been introduced in this conceptually driven manuscript.

Acknowledgments

We thank Dr. Graciela Rosemblat for her assistance and helpful comments. This work was supported by the Intramural Research Program of the NIH, National Library of Medicine.

References

[1] Napoli PM. Consumer use of medical information from electronic to paper media: A literature review. In: Rice RE, Katz JE, (eds). The internet and health communication: experiences and expectations. Thousand Oaks: Sage, 2001:79-98.

[2] Dutta Bergman M. Theory and practice in health communication campaigns. Health Communication 2005 18(2):103-122.

[3] Baker LM, Pettigrew KE. Theories for practitioners: two frameworks for studying consumer health information-seeking behavior. Bull Med Libr Assoc. 1999 Oct;87(4):444-50.

[4] Dervin B. Useful theory for librarianship: Communication, not information. Drexel Library Quarterly 1977;13(3):16-32.

[5] Belkin NJ. Anomalous states of knowledge as a basis for information retrieval. Canadian Journal of Information Science 1980;5:133-143.

[6] Kuhlthau C. Longitudinal case studies of the information search process of users in libraries. Library and Information Science Research 1988;10:257-304.

[7] Marchionini G. Information seeking in electronic environments. Cambridge UK: Cambridge University Press. 1995.

[8] Zeng QT, Crowell J, Plovnick RM et al. Assisting consumer health information retrieval with query recommendations. J Am Med Inform Assoc. 2006 Jan-Feb;13(1):80-90.

[9] McGuire WJ. Input and output variables currently promising for constructing persuasive communications. In: Rice RE, Atkin CK (eds). Public communication campaigns. Thousand Oaks: Sage, 2001: 22-48.

[10] Wimmer RD, Dominick JR. Mass communication research: an introduction. Belmont, CA 2006 Wadsworth, 8th edition.

[11] Capella JN. Integrating message effects and behavioral change theories: organizing comments and unanswered questions. Journal of Communication 2006 56:S265-S279.

[12] Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res. 2001 Dec;16(6):671-92.

[13] Greenberg L, D'Andrea G, Lorence D. Setting the public agenda for online public health: a white paper and action agenda. Washington, DC: URAC [Utilization Review Acc reditation Commission], 2003.

[14] Mitchell JA, Fun J, McCray AT. Design of the Genetics Home Reference: a new NLM consumer health resource. J Am Med Inform Assoc. 2004 Nov-Dec;11(6):439-47.

[15] Logan RA, Fun J, Cheh M. User perceptions of Genetics Home Reference: a study of a health information Web site's image. Intl Comm Assoc Ann Meet. 2006 June.

[16] van der Rijjt GAJ. Determinants of the consumption of health information in the media. Eur J Commun Res. 1998;23(3):255-69.

[17] Bakker TA, Ryce AN, Logan RA, Tse T, Hutcherson L. A consumer health informatics (CHI) toolbox: challenges and implications. Proc AMIA Symp 2005: 21-5.

Address for correspondence

Robert A. Logan, Ph.D.,
Lister Hill National Center for Biomedical Communications,
Building 38A. Bethesda, MD USA. 20894.
Email: logan@nlm.nih.gov

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

A Japanese Model of Disease Management

Naoki Nakashima[a], Kunihisa Kobayashi[b], Toyoshi Inoguchi[b], Daisuke Nishida[c], Naomi Tanaka[c], Hiromi Nakazono[d], Akihiko Hoshino[e], Hidehisa Soejima[e], Ryoichi Takayanagi[b], Hajime Nawata[d]

[a] *Department of Medical Informatics, Kyushu University Hospital, Japan*
[b] *Department of Medical and Bioregulatory Science, Graduate School of Sciences, Kyushu University, Japan*
[c] *Carna Health Support, Japan*
[d] *Graduate School of Medical Sciences, Kyushu University, Japan*
[e] *Saiseikai Kumamoto Hospital, Japan*

Abstract

We started a disease management model, Carna, that includes two programs: one for primary prevention of life-style diseases and one for secondary/tertiary prevention of diabetes mellitus. These programs support the family doctor system and education for participants to allow the concept of disease management to take root in Japan. We developed a critical pathway system that can optimize health care of individual participants by matching individual status. This is the core technology of the project. Under the primary prevention program, we can perform the health check-up/ instruction tasks in the 'Tokutei Kenshin', which will start for all Japanese citizens aged 40 – 74 years in April 2008. In the diabetic program, Carna matches doctors and new patients, prevents patient dropout, supports detection of early-stage complications by distributing questionnaires periodically, and facilitates medical specialists' cooperation with family doctors. Carna promotes periodic medical examinations and quickly provides the result of blood tests to patients. We are conducting a study to assess the medical outcomes and business model. The study will continue until the end of 2007.

Keywords:

disease management, life style disease, diabetes mellitus, critical pathway

Introduction

Disease management is a system of coordinated health care interventions and communications for populations with conditions in which patient self-care efforts are significant [1]. Secondary and tertiary prevention of specific diseases is easier than primary intervention because we cannot predict who will be affected by certain diseases. Therefore, secondary and tertiary prevention is already practiced in the United States [2]. In Japan, however, primary prevention has predominated. This is because the medical institutions and patients have not been adequately motivated to participate in disease management. Medical institutions are independent of insurers, and all citizens of Japan are insured. We have developed a Japanese model of disease management, Carna, which has focused on diabetes mellitus since 2005.

In June 2006, the Japanese Government amended the Medical Care Law to establish a particular health check-up program that includes health instruction (Tokutei-Kenshin). The program will begin in April 2008 [3]. This will be a possible trigger to spread disease management in Japan.

Here, we first introduce the Tokutei-Kenshin system. Then we outline our disease management system and describe the issues involved in system development.

Materials and methods

The Tokutei-Kenshin system

In June 2006, the Japanese Government amended the Medical Care Law to establish an annual health check-up/ management system for all citizens aged 40–74 years. This system will greatly affect insurers because it will cover 57 million citizens (45% of entire Japanese population) and will involve all insurers from April 2008. Insurers will be penalized economically from 2013 when stated goals are not achieved. All Japanese people have basic health care insurance. Therefore, this law will be a very important and have a significant influence on health management in Japan. However, it remains unclear who will provide what kinds of services, even though the government has provided a basic framework for the system.

The government has often told insurers and citizens about the significance of measures against lifestyle disease, and about the Tokutei-Kenshin after the law was passed. Covered citizens can refuse to receive a health-check up. Some insurers are complaining, although the law has been passed, as they are the responsible entity. The health check-up rate will increase by degrees.

The annual health check-up

The annual health check-up involves:

1. Questionnaire (enquiry pertaining to weight change, smoking, exercise).

2. Physical examination (height, weight, body mass index, waist, blood pressure).
3. Blood chemistry tests (triglyceride, HDL-cholesterol, LDL-cholesterol, GOT, GPT, -GTP, Cre, blood glucose (fasting or postprandial), HbA1c, uric acid).

Stratification involves two steps. On the basis of results, participants are assigned to one of three risk groups. The waist and body mass index determine the first assignment. Risk factors identified by blood chemistry tests and smoking history are tallied, and participants are assigned to the 'information provided groug', 'motivation support group', or the 'aggressive support group' (Figure 1).

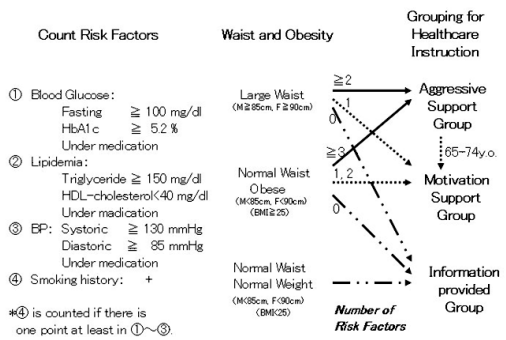


Figure 1 - Stratification into risk groups according to the results of physical examination and blood chemistry tests.

Insurers need to provide health care instruction once to individuals assigned to the motivation support group, and repeatedly to individuals assigned to the aggressive support group. The information provided group is not given health instruction. A physician, a health nurse or a registered dietitian should provide the first health care instruction each year. Non-face-to-face instruction via telephone or information and communication technology (ICT) system is allowed for additional instruction for the aggressive support group. Insurers do not need to provide health care instruction to the person who is under medication for lifestyle disease when they are classed in the motivation support group or the aggressive support group.

If the results of the physical or blood examination are out of the normal clinical range described below, the insurer must encourage the insured person to visit a clinic and confirm the visits that occurred each year.

1) Blood glucose

Fasting	≥	126 mg/dl	or
HbA1c	≥	6.1%	

2) Blood lipids

Triglyceride	≥	300 mg/dl	or
HDL-cholesterol	<	35 mg/dl	

3) Blood pressure

Systolic	≥	140 mmHg	or
Diastolic	≥	90 mg/dl	

4) LDL-cholesterol ≥ 140 mg/dl

All data obtained from the Tokutei-Kenshin will be digitized in a standard protocol as HL7 CDA, and in standard cords as JLAC 10. Insurers will have to maintain records for each insured person as long as they remain insured. The records will contain health check-up and instruction data. If the individual changes insurer as they change occupation, the former insurers will transfer the data to the new insurer so there is no gap in the record. With these systems in place, it is anticipated that the Japanese Government will have access to huge amounts of anonymous data that can be used for statistical purposes. Insurers can keep the data without anonymising it because, for example, they have to analyze it with the data from the medical institutions. Insurers will be able to store data more effectively when the online reimbursement project is fully realized in 2011.

The Japanese Government suggests that the insurer pay an additional 10% contribution to the medical costs of those aged 75 year or over if it does not achieve the stated goals, or be rewarded by a 10% discount if stated goals are achieved from 2013. The Japanese Government is asking for a 25% decrease in the number of diabetes and pre-diabetes patients per insurer, for example.

Insurers will be allowed to outsource the required tasks to health care provider companies. Thus, many tasks will be outsourced to Internet data centers.

Initially, insurers will engage a health care provider company for outsourcing. After registration of the covered citizens, the health care provider will begin service. If providers do not produce adequate outcomes, insurers must pay the penalty as they must take overall responsibility.

Strategy for development of a Japanese model of disease management

We first focused on secondary/tertiary prevention of diabetes mellitus as a Japanese model of disease management, Carna. We then developed a program for the prevention of diabetes complications. After the Japanese Government decided upon the Tokutei-Kenshin system, we developed a program for lifestyle improvements. Our program allows insurers to outsource the Tokutei-Kenshin tasks.

Our disease management program is based on an outbound-call center that manages all information obtained from insurers, medical institutions and individual participants. The call center staff (a nurse or a dietitian) uses the telephone, regular mail and e-mail according to their ICT literacy. The call center always functions as the center of information in data management and communication. We store all information securely at an Internet data center.

We collect information on an individual's initial/improved lifestyle (food, exercise, smoking, alcohol, stress, daily sleep, so on) and knowledge about lifestyle diseases. We also collect the result of blood tests and physical check ups (waist, weight, height, blood pressure, and so on). For primary prevention, the Tokutei-Kenshin includes an initial questionnaire-based interview pertaining to lifestyle. In follow-up health instruction for the aggressive support group, we mainly use e-mail by call center staff. In second-

ary/tertiary prevention, call center staff also communicate with patients by telephone or regular mail.

The main target outcome of the primary prevention program is a decrease in the prevalence of pre-diabetes and diabetes. The main target outcome of the secondary/tertiary prevention program is a decrease in the number/degree of diabetic complications. We collect information about these outcomes on a regular basis.

Insurers will be clients of the primary prevention program after the Tokutei-Kenshin starts in 2008. The family doctors/medial institutions will be clients of the secondary/tertiary prevention program. The Japanese medical insurance pays the lifestyle instruction fee to the medical institutions. The model suggests that the outsourcing be paid for from the fees.

We developed the project with high regard for (1) adaptation to the Japanese medical system, (2) quality control, (3) appropriate matching of services to individuals, (4) adaptation to the Japanese political direction in Japan, (5) efficient and secure data management, and (6) ethical considerations.

(1) Adaptation to the Japanese medical system

Medical institutions and citizens in Japan are less motivated than their counterparts in the United States to participate in disease management. Medical institutions are independent of insurers because all Japanese citizens are covered by public medical insurance and allowed access to any medical institutions. Thus, we have placed importance on providing incentives to medical institutions and to citizens, and we call this a 'Japanese model' of disease management.

(2) Quality control

As disease management will involve call center staff who must deal with enormous numbers of participants and medical institutions, we have given much thought to the quality control of services by compiling tasks and developing ICT systems in the call center. Quality control is particularly challenging because of the scale of the new system.

(3) Appropriate matching of services to individuals

It is difficult to manage and improve the disease status, lifestyle and self-care of individuals by uniform instruction/intervention because individuals vary considerably in basic character, family lifestyle, psychology and health/disease status. Thus, we have attempted to personalize the match-up between services and individuals.

(4) Adaptation to the political direction in Japan

The Japanese Government is focusing on preventing lifestyle diseases. The Tokutei-Kenshin is a product of the new policy. It has been necessary to watch the government and we need to adapt our activities to its policies if we expect the disease management business to take root in Japan.

(5) Efficient and secure data management

For ordinary disease management, data management including collection, storage, analysis and feedback are essential. The scale of the Tokutei-Kenshin system means we expect and are prepared to deal with huge amount of standardized data.

(6) Ethical considerations

There are many legal problems and a new social framework is needed for disease management providers who ensure cooperative, convergent and seamless service to patients by multiple medical institutions and health care service providers. We adhere to the Japanese privacy protection law. In addition, we established a new framework for disease management providers in protecting the privacy of patients' health information, and we propose appropriate handling of patients' health information by multiple service providers working together as a disease management consortium.

Results

(1) Strategy to allow disease management to take root in Japan

To motivate clinics and patients to participate in disease management, we support the family doctor system and education for patients.

In clinics, we match doctors and new patients when they are affected with a lifestyle disease. We also attempt to prevent patient dropout by telephone contact, support patient education and detection of early-stage complications by means of periodic questionnaires, and facilitate medical specialists' cooperation according to the timing described in the critical pathway system.

For patients, we promote medical care described in the critical pathway, report the results of blood tests quickly, and provide 'Carna points' as rewards for the patient's efforts (for instance, regular clinic visits) and for improvement in their diabetic condition (HbA1c). We exchange the points for coupons with which they can obtain certified health-related products such as healthy foods and exercise goods.

(2) Critical pathways for quality control for appropriate matching of services to individuals, and for adaptation to the political direction in Japan

We developed a region-related, outcome-oriented critical pathway as the core competency in the call center. We also standardized workflow in the call center calling 'algorithm'. The critical pathways and the parts of the algorithm are digitized. We prepared an education system with structured questionnaires and comprehensive teaching materials that are closely related to the personalized critical pathway.

We had two kinds of outcome-oriented critical pathways by the end of 2006. One is for the primary prevention program (lifestyle improvement program); the other is for the secondary/tertiary prevention program for diabetes mellitus.

Features of the critical pathway for lifestyle improvement program edge are:

- Using five kinds of critical pathway matching each stage of Prochaska stage model (pre-contemplation

stage, contemplation stage, preparation stage, action stage, maintenance stage) [4].

- Matching the framework of the Tokutei-Kenshin (see Figure 2).

Features of the critical pathway for diabetes mellitus (secondary/tertiary prevention) that we have developed [5] are:

- Scheduling of medical services based on clinical guidelines produced by the Japanese Diabetic Society.
- Supporting general care of diabetic outpatients including timely reminders of the need to visit medical specialists such as an ophthalmologist and a diabetologist.
- Using 'the overlay method' to create an optimal personalized critical pathway for each patient. A personalized critical pathway is created by overlay with a basic sheet for regular examination, and optional sheets matching patient's treatments, the severity of the diabetic complications, and the patient's level of knowledge. We can create 2880 different of critical pathways [5].
- Modifying continuously as the patient's condition

(3) Personalized communication based-on patient characteristics

Patients' responses to interventions vary because the characters of patients vary. For successful intervention, we determine the patient's character type during the registration process, and we depend on this information to personalize our communication with the patient. This approach may also decrease the call center staff's stress.

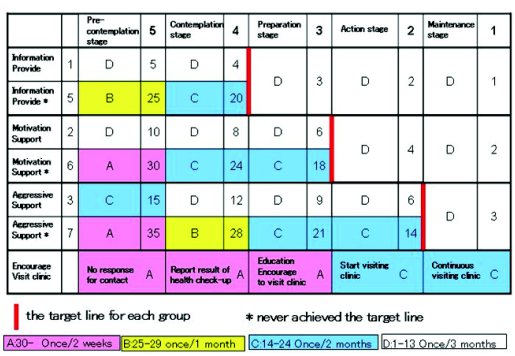


Figure 2 – Algorithm for telephone call frequency based on the Prochaska stage model and the stratification of the Tokutei-Kenshin. Each risk group has a target line on the staging model

(4) Other algorithms and ICT system for efficient and secure data management

We developed another algorithms ICT system as shown in Figure 3. We used an application service provider system for ICT to input participant records.

We are using an Internet data center in Fukuoka city, Japan for database servers. We send data over the Internet via a virtual private network from the call center.

In the primary prevention program, the initial health instruction requires a face-to-face meeting. We tried a teleconference system using the Internet as a non-face-to-face method. We used VIPS teleconference system developed by Kyuden Infocom Co., and it worked without any problems. In the future, we want to use teleconferencing for the initial health instruction because there are many people living in the countryside in Japan who have access to the Internet. This field is fit to use telemedicine [6]. However, the new law does not allow non-face-to-face methods in the initial instruction.

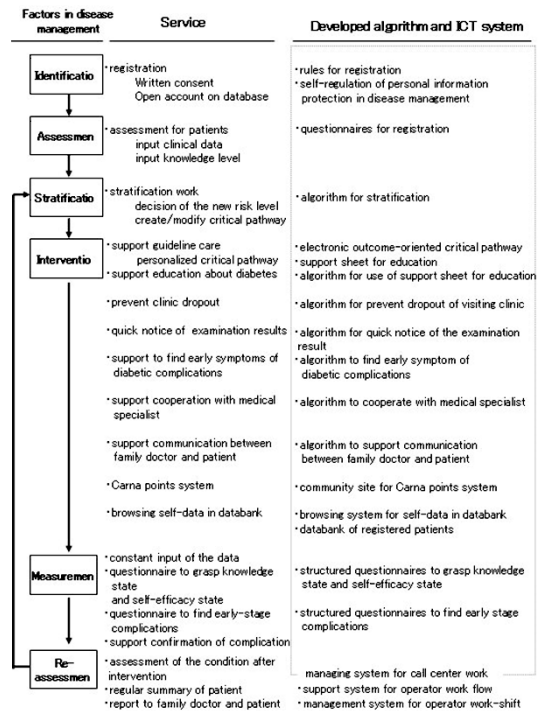


Figure 3 – Disease management services and rules, algorithms, and ICT system in the secondary/tertiary prevention program for diabetes mellitus

(5) Ethical considerations

We quantified the risk of disclosure in terms of information value, the threat arising from inadequate ICT security and areas of vulnerability and showed that the highest risk was posed by databases containing individual patient profiles. Consequently, we need regulations pertaining to the provision of health information and general classifications for various types of patient information that we deal with in disease management work.

We are currently conducting a study to assess the medical outcomes and the business model. The study will be completed at the end of 2007. The ethics committee of the Graduate School of Medical Sciences, Kyushu University approved this study.

Discussion

The Disease Management Association of America defines disease management as a system of coordinated health care interventions and communications for populations with conditions for which patient self-care efforts are significant, and full-service disease management includes the six components: (1) population identification processes, (2) evidence-based practice guidelines, (3) collaborative practice models to include physician and support-service providers, (4) patient self-management education (may include primary prevention, behavior modification programs, and compliance/surveillance), (5) process and outcomes measurement, evaluation and management, and (6) routine reporting/feedback loop (may include communication with patient, physician, health plan and ancillary providers, and practice profiling).

Our disease management 'Carna' meets the definition and includes all six components.

The Tokutei-Kenshin system comes close to meeting the definition of disease management because the Japanese Government focused on metabolic syndrome for the design of the Tokutei-Kenshin system. Thus, we can say that a nationwide disease management project will begin in 2008, although the government has left the details to the insurers who will undertake it. However, typical Japanese insurers are not familiar with health check-ups that are followed by stratification and health care instruction, and they do not have practical knowledge in disease management. Thus, many of them will outsource the task to disease management providers. We fear the possibility that outsourcing will be awarded to inferior providers, and believe the Japanese Government should provide strict recommendations pertaining to the skills of providers who accept outsourcing of the Tokutei-Kenshin tasks.

In the future, we will add a program aimed at chronic disease using the same strategy.

Acknowledgments

This study was supported in 2005 and 2006 by 'a fund to create a new health care service' in the Japanese Ministry of Economy, Trade and Industry. We sincerely appreciate Kyuden Infocom Co., Kyushu Electric Power Co., Tokyo Marine Nichido Co., and Fukuoka Prefecture Medical Association for their support of this study.

References

[1] Disease Management Association of America: Definition of Disease Management, 2003.1.6., available from: http://www.dmaa.org/dm_definition.asp

[2] Rubin RJ, Dietrich KA, Hawk AD. Clinical and economic impact of implementing a comprehensive diabetes management program in managed care. J Clin Endocrinol Metab. 83: 2635-2642, 1998

[3] Japanese Ministry of Health, Labour and Welfare: Particular health check-up/instruction, available from: http://www.mhlw.go.jp/bunya/kenkou/seikatsu/pdf/02.pdf

[4] Ruggiero L, Rossi JS, Prochaska JO, Glasgow RE, de Groot M, Dryfoos JM, Reed GR, Orleans CT, Prokhorov AV, Kelly K. Smoking and diabetes: readiness for change and provider advice. Addict Behav. 24:573-578, 1999

[5] Kobayashi K, Nakashima N, Inoguchi T, Nishida D, Tanaka N, Hoshino A, Hamada M, Kozuma Y, Nakaguma H, Matsushita R, Soejima H, Takayanagi R, Nawata H. Development of a new critical pathway for regional medical network for outpatients with diabetes mellitus. J. Japan Diab. Soc. 49: 817 – 824, 2006

[6] Garcia-Lizana F, Sarria-Santamera A. New technologies for chronic disease management and control: a systematic review. J Telemed Telecare. 13: 62-68, 200.

Address for correspondence:

Naoki Nakashima
Maidashi 3-1-1,
Higashi-ku,
Fukuoka 812-8582,
Japan
E-mail address: nnaoki@info.med.kyushu-u.ac.jp

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Towards Sustainability of Health Information Systems: How Can We Define, Measure and Achieve It?

Sebastian Garde[a, b], Carola M. Hullin[a, b], Rong Chen[c], Thilo Schuler[d], Jana Gränz[a, e], Petra Knaup[f], Evelyn J.S. Hovenga[a]

[a] *Health Informatics Research Group, Central Queensland University, Melbourne VIC & Rockhampton, QLD, Australia*
[b] *Austin Centre for Applied Clinical Informatics, Austin Health, Heidelberg VIC, Australia*
[c] *Department of Biomedical Engineering, Linköping University, Sweden*
[d] *Department of Medical Informatics, University of Freiburg, Germany*
[e] *Faculty of Computer Science, University of Applied Sciences Ulm, Germany*
[f] *Department of Medical Informatics, University of Heidelberg, Germany*

Abstract

Health information systems (HIS) in their current form are rarely sustainable. In order to sustain our health information systems and with it our health systems, we need to focus on defining and maintaining sustainable Health Information System building blocks or components. These components need to be easily updatable when clinical knowledge (or anything else) changes, easily adaptable when business requirements or processes change, and easily exchangeable when technology advances. One major prerequisite for this is that we need to be able to define and measure sustainability, so that it can become one of the major business drivers in HIS development. Therefore, this paper analyses general definitions and indicators for sustainability, and analyses their applicability to HIS. We find that general 'Emergy analysis' is one possibility to measure sustainability for HIS. Based on this, we investigate major enablers and inhibitors to sustainability in a high-level framework consisting of four pillars: clinical, technical, socio-technical, and political/business.

Keywords:

health information systems, electronic health records, sustainability, *open*EHR, computerized medical record systems

Introduction

Information, knowledge management and communication technologies are crucial enablers of system change that can play a vital role in substantially transforming healthcare systems and prevent their failure[1]. From this perspective, we argue further that the sustainability of our health systems depends largely on the sustainability of our Health Information Systems (HIS).

The past few years have seen a myriad of developments and deployments of HIS. Some have very limited focus, other operate on a regional scale, and a few initiatives are underway to establish nationwide HIS, for example in the form of Shared Electronic Health Record (EHR) Systems. The World Health Organisation (WHO) requests international collaboration [1] - good examples include the *open*EHR (http://www.openEHR.org), HL7 (http://www.hl7.org), and the joint Detailed Clinical Models (DCM, http://detailedclinicalmodels.org) initiatives.

With patients being increasingly mobile and treatments and health care providers increasingly specialized, interoperability of HIS has become critical for sustaining current processes. Patient data can be relevant for over 100 years, thus sustainability of patient data is critically important. Not only will sustainability and interoperability save money, we also expect a significant positive clinical impact (cp. e.g. [2]).

While defining HIS failure and success is complex, and current evidence on HIS success and failure rates is insufficient, the "*best current estimate is that HIS failure is an important problem*" [3]. If current customs prevail, very few of these systems will be sustainable, let alone semantically interoperable, costing lives and money. According to Haux, a lot of research and application is necessary to further develop and investigate HIS architectures and infrastructures, in order to identify sustainable approaches [4]. In this context, the aim of this paper is to

- Review general definitions and measurements of sustainability
- Analyse their applicability to HIS
- Investigate major inhibitors and enablers of sustainability in a high-level framework for building sustainable HIS. In this framework, we will relate to the *open*EHR architecture as one candidate for an approach to sustainable development of HIS.

Materials and methods

We conducted a literature review on 'sustainability', a summary of which is presented in this paper. In addition to being heavily involved in many of the following, we

1 Cp. Medinfo 2007 Call for Papers.

reviewed the literature and web resources available on socio-technical issues, Electronic Health Records (EHR), failures of Health Information Systems, recent reports from national EHR initiatives, *open*EHR (http://www.openEHR.org, [5]) as an advanced architecture for EHRs and interoperability, Interoperability Frameworks like the one provided by NEHTA [6], and initiatives like the Detailed Clinical Models initiative.

The enabling/inhibiting pillars for sustainability described in in the second part of this paper are derived from common threads found in the Health Informatics literature on successful and failed systems implementations as well as own experiences. These pillars are related to the structure of Interoperability Frameworks like [6].

Results

Definition and metrics for sustainability

It is believed that the word *sustainability* (German: *Nachhaltigkeit*) was used for the first time in 1712 by the German forester and scientist Hans Carl von Gilinscee in his book *Sylvicultura Oeconomica* [7]. While according to the Canadian *Sustainability Now* initiative more than 300 definitions of sustainability exist (http://www.sustainability.ca), the probably best-known definition stems from the World Council on Environment and Development [8]. It defines sustainable development as that which "*…meets the needs of the present without compromising the ability of future generations to meet their own needs*". From the authors' perspective, the simplest, most generic and compelling definition is found on Wikipedia: "*Sustainability: the ability to continue a defined behavior indefinitely".* However, this simplicity is spoilt, once more specific definitions are required – given that the "the term itself is being applied to so wide a range of issues that it can no longer retain only a single meaning" [9]. If this is the case, how do we know that we have achieved sustainability? In 2003, Maine brought attention to this lack of quantitative indicators of sustainability [9] and a decade earlier the International Institute for Environment and Development ([10], p.2) had already concluded that "*the need for sustainability analysis and particularly for indicators of sustainability is a key requirement to implement and monitor the development of national sustainable development plans […]*".

To quantify sustainability, Maine suggests "*narrowing the use of the term strictly to physical processes, for this appears the only way to achieve the establishment of a secure and robust metric of sustainability […]*" [9]. Consequently, he argues for a "*rigorous metric of sustainability derived from basic scientific principles, and avoiding the application of the term to sociological issues such as the longevity of an organization or society.*" He proposes the energy of reclamation of all outputs of anthropogenic (=derived from human activities) processes as a metric for sustainability. An important attribute any system requires to be sustainable would then be the minimal production of 'waste' or "*the amount of energy that is NOT used to reclaim waste*" [9].

A series of sustainability indicators is based on 'emergy synthesis' as introduced by Odum in [11]. Emergy (with 'm', not 'n') is an abbreviation of the term "embodied energy". Without going into complex mathematical definitions of emergy, it shall be said that emergy expresses the cost of a process or a product in solar energy equivalents, which is regarded as the ultimate energy source. Odum's innovation established a medium for environmental accounting that for the first time made it possible to express economic commodities, services, and environmental work of all kinds on a common basis as emergy [12]. In other words, by expressing the value of products in emergy units (*emjoule*), it becomes possible to compare 'apples and pears' [13].

Once the total number of input flows into a system has been identified and based on this the total emergy driving a process has been evaluated, a set of indicators can be calculated to illuminate different aspects of sustainability as the following important indicator developed by Brown and Ulgiati ([14]):

$$Sustainability\,Index = \frac{Emergy\,Yield\,Ratio}{Environmental\,Loading\,Ratio} = \frac{\frac{Y}{F}}{\frac{N+F}{R}} \quad (1)$$

This index is also called the "Emergy Sustainability Index" (ESI). The *Emergy Yield Ratio* is defined as the ratio of the emergy of the output of the system (Y) and the emergy of purchased services and resources that are input to the system (F). *Emergy Loading Ratio* is defined as the sum of the emergy of local non-renewable sources (N) and purchased resources/services (F) divided by the emergy of the free environmental emergy available from local renewable sources (R).

More recently a joint initiative of Yale and Columbia University, in collaboration with the World Economic Forum and the Joint Research Centre of the European Commission constructed an Environmental Sustainability Index (also called ESI, http://sedac.ciesin.columbia.edu/es/esi/index.html) and compared it to other sustainability indicators such as the Ecological Footprint Index measuring the area of productive land and water appropriated exclusively to produce the resource used and to assimilate the waste generated [15]. Zhao and colleagues further introduce a modified form of ecological footprint calculation by combining emergy analysis with conventional ecological footprint analysis [15].

Sustainability in health and health information systems

Given the long history of sustainability reaching back into the 18th century, it is astonishing that there are no agreed definitions for sustainability of health systems or health information systems – clearly the generic definitions are not sufficient for any measurement and the more specific definitions and indicators of sustainability are not applicable without restrictions to the area of health. Nonetheless we often argue that our health systems are not sustainable as for example Enrico Coiera in a recent paper: "*The health system at present is one that consumes enormous resource, and generates enormous waste, and* ***would not meet any criterion of sustainability****. Injecting new interventions from 'outside' the system, as we currently do in health informatics, is itself not a sustainable approach, as*

the capacity for external designers to meet all the evolving needs of those inside will just never be there" [16]. Thus, intuitively we know that our health information system infrastructure as a whole is not sustainable – however we cannot measure it and thus are not able to take systematic corrective action.

In measuring sustainability of a health information infrastructure, we believe that in analogy to environmental sustainability, it is necessary to 'compare apples and pears' and thus have a common unit like emergy that achieves this for us. In essence, we need to consider what the inputs and outputs and storages of the system under investigation are. We then need to analyse - similar to environmental emergy analysis - which parts are renewable or non-renewable or used non-renewably. Where previously unknown (because no previous studies exist), we need to determine what factors we can use to convert these inputs and outputs into emergy units as detailed by Odum in [11]. Once we have achieved this, all the environmental sustainability indicators that are based on emergy can be applied to health information systems as well. For a given system, we need to identify and analyse the inputs, outputs and stored resources of the system according to Table 1. Other indicators like the ecological footprint seem to be less applicable for HIS.

In contrast to environmental sustainability, the differences between renewable and non-renewable sources however are not always that explicit - e.g. the labour of skilled workers are not always simply renewable – but can be very hard to come by. However, proper training/education programs can make a difference in the long run. For environmental sustainability these resources would be classified as 'used non-renewably' without further distinction [12]. For the analysis of HIS infrastructures, this may not be sufficient and we are currently investigating the use of a 'renewability factor' to rectify this.

Inputs, outputs, storages for health information systems and inhibitors for their sustainability

Table 1 – Analysis of items that are input or output of the system or are stored within the system.

Item	Input/output of the system and any items that stored within it. Some items are the same as typically used for environmental emergy analysis, but others differ.
Data	Raw data measured in joules, grams, dollars or any other appropriate unit.
Solar Emergy per Unit	Factor to transform the data into solar emergy.
Solar Emergy	Calculated: Data x Solar Emergy per Unit

In the following, we investigate these aspects – and especially the enablers and inhibitors for sustainable Health information systems in a high-level framework based on four pillars (each consisting of several high-level building blocks):

- clinical,
- technical,
- socio-technical,
- political & business.

These pillars are related to the structure of Interoperability Frameworks like [6]. Our framework is intended to 'get it right' on a high-level, not about providing all the details – as these details are (with a few exceptions mentioned in the following) relatively well researched.

Some of the inputs and outputs identified are similar to those investigated for environmental sustainability, however some are quite different. This is largely due to the fact that knowledge of various kinds can be seen as one of the most precious resources for us. Table 2 summarises some of the inputs, outputs, and stored items typical for a health information infrastructure, which are untypical from an environmental sustainability point of view and identifies typical inhibitors for sustainability with regard to each item.

Pillar 1: Clinical building blocks

The clinically most important building block for a clinical system is the agreement on clinical content. This fosters semantic interoperability between systems and provides clear meaning – so that we can exchange and migrate data between different systems and support clinical decision making. This clinical domain knowledge needs to be managed and maintained – a complicated task that eventually has become feasible ([5]), although it will always remain difficult on a national or international scale to reach agreement. The separation of technical and clinical concern through *open*EHR's two-level-modelling paradigm seems to be well suited to enable this because it clearly separates clinical content definition from technical concerns.

We suggest the development and international use of a repository of clinical content models that are freely available so that 'flexible standardisation' of this content can occur[2]. For example, *open*EHR archetypes are particularly well suited to serve as a standard form for these clinical content models because archetypes are intuitive to clinicians, but also formal specifications of clinical content technicians can work with.

Pillar 2: Technical building blocks

The technology chosen must be able to cope with the constant changes of health care and health care knowledge without having to change enormous amount of source code (and wait for the vendor to implement it). It must provide the technical basis for semantic interoperability and sustainability. This is not only important because of more and more specialised providers and more and more mobile patients, it also enables the migration of systems without losing considerable amounts of patient data, thus also avoiding vendor lock-in.

2 See http://www.archetypes.com.au for the Archetype Finder, which is designed to support this task as well as the open source Java implementations of *open*EHR at http://www.openehr.org, developed by some of the authors of this paper.

To achieve semantic interoperability, "*[y]ou need the ontology, the information model and services. [..] If you have one and don't have the others, it won't help*" [17]. It is the author's view that ontology and services are relatively well understood – the information model however is largely ignored. This is where an approach like the *open*EHR approach, which is based on a stable and generic information model is of major importance. This works similar to the Java programming language, which decouples itself from the operating system by translating the Java source code to Java bytecode. This code is then run on a native Java virtual machine – thus enabling portability. In a similar way, the *open*EHR two-levelling modelling decouples the technical knowledge (the information model of the software) from the clinical knowledge (expressed in archetypes) to achieve semantic interoperability.

Most importantly, technology needs to be designed to consist of largely independent components, so that replacement can occur without endangering the sustainability of the infrastructure as a whole. Open source implementations will help to validate, improve, evolve their specifications and educate the early implementers [18]. Moreover, if the initial code base is good enough for others to collaborate there is no need to re-invent wheels and thus open source components can serve as building blocks for high level HIS applications. Linux, Apache, OpenBSD, JBoss, Hibernate, Ant and many more are all part of the backbone of our current technical infrastructure and they contribute enormously to sustainability.

Pillar 3: Socio-technical building blocks

While acknowledging the utmost importance of socio-technical issues ranging from comprehensive change management, proper localisation of clinical systems, sufficient training, etc., this has in theory been very well investigated (although not often enough implemented in practice), and "*[s]ocio-technical systems (STS) analysis has provided us with a powerful framework with which to analyse the reasons behind the poor acceptability, uptake and performance […]*" [19]. We therefore refrain here from elaborating on this topic.

Pillar 4: Political/ business building blocks

No matter how much sense this framework makes on a technical, clinical or socio-technical level – sustainability (as well as interoperability) needs to become a major business driver to become reality! We need to ensure that politics is informed and business drivers are 'right'. For this sustainability needs to be measured and a case made for public sector and regional healthcare information systems to be based on open source software to remove the risk associated with any given vendor.

In an era of constant change, political/business decisions do not often hold up long enough to see the rewards of a decision of e.g. implementing a nationwide EHR – as this commonly takes years from planning to roll-out. For this

Table 2 - Examples for inputs, outputs, and stored items in a health information infrastructure – and inhibitors that currently often prevent their sustainability

Sustainability Items	**Description**	**Examples for inhibitors to sustainability**
Health Informatics Knowledge and Skills	Knowledge and Skills of Health Informatitions/ Health IT and IS Professionals. This is vital as we need to work with limited resources to fulfill the great demand.	• Starting similar initiatives from the beginning over and over again, thus losing knowledge inherent in unsustainable systems, which then needs to be recreated • Socio-technical issues, including localisation issues and insufficient change management • Political change and infrastructure not set up appropriately for political change. Independent umbrella organisations may be one solution. • Wrong infrastructure to cope with major changes (e.g. disaster management)
Clinical/ Patient Information	The clinical information on a patient stored in a clinical system. If this information is lost, it has major consequence for patient life and money. We need to sustain clinical data for 100years and more.	• Migration of systems without loss of patient data not feasible (e.g. information model not sufficiently clear) • Vendor lock-in • Insufficient system ability to match clinical practice / workflow • Wrong business drivers • No open source or otherwise available and agreed (e.g. standardised) specifications for data that needs to be shared
Clinical knowledge	The clinical knowledge of Health Professionals – maintained, structured and evidence-based.	• Hard-coded clinical knowledge • Evolvement of clinical practice and general processes which causes systems to slowly become obsolete

reason, we propose (independent) umbrella organisations on a national and eventually internationally level that render political decisions more predictable and sustainable. These organizations can provide leadership and stability for specific purposes.

Finally, while obviously appropriate funding is essential, it does not seem to be wise to provide more and more funding to a system that struggles - without identifying and addressing the fundamental problems within the 4 pillars.

Discussion and conclusion

As shown in this paper, there are many problems to solve in achieving any state, which could be called sustainable. Some are technological, some are socio-political and some organizational. However, most importantly, the lack of a definition for sustainability and agreed and standardised metrics to measure sustainability in health is problematic because without we cannot quantify the status quo of sustainability at any given point of time - and as a consequence sustainability will be largely ignored by decision makers as an important business driver. Indicators generally simplify in order to make complex phenomena quantifiable so that information can be communicated efficiently to decision makers. We thus have to ensure that suitable indicators are applied to make the complex phenomenon sustainability quantifiable for decision makers. In this paper, we showed one possible way of developing such an indicator.

Shabo's Model for the Sustainability of Longitudinal EHRs [20] is one that tackles the concrete problem of sustaining EHRs. As such Shabo's model is more specific to this concrete problem, but also limited to it; his suggestions are in harmony with our results.

While Return on Investment (ROI) is an inherent part of this this paper, its quantification is not part of this paper. For a first estimate, a similar model as the model used by Walker and colleagues to quantify the value of health care information exchange and interoperability [2] could be employed. Also, the World Business Council for Sustainable Development, has formulated the business case for sustainable development and argues that "*sustainable development is good for business and business is good for sustainable development*". The same we believe is true is for the sustainable development of HIS. We need to develop such a business case for sustainability for our health information infrastructure and change business practices on every level. One important step for this is to develop, evaluate and use indicators for sustainability that can be used by decision makers to quantify sustainability to justify expenditures on fighting barriers to sustainability and take a systematic approach towards sustainability of HIS. This can be based on the analysis presented in this paper.

References

[1] WHO. WHA58.28 - Resolution on eHealth. http://www.who.int/gb/ebwha/pdf_files/WHA58/WHA58_28-en.pdf. 2005 (Last access: 08.03.2007).

[2] Walker J, Pan E, Johnston D, Adler-Milstein J, Bates DW, Middleton B. The value of health care information exchange and interoperability. Health Aff (Millwood) 2005: Suppl Web Exclusives: W5-10-W15-18.

[3] Heeks R. Health information systems: failure, success and improvisation. Int J Med Inform 2006: 75 (2): 125-37.

[4] Haux R. Individualization, globalization and health - about sustainable information technologies and the aim of medical informatics. Int J Med Inform 2006: 75 (12): 795-808.

[5] Garde S, Knaup P, Hovenga E, Heard S. Towards Semantic Interoperability for Electronic Health Records: Domain Knowledge Governance for *open*EHR Archetypes. Accepted for Methods of Information in Medicine 2007.

[6] National E-Health Transition Authority (NEHTA). Interoperability Framework. 2006.

[7] von Carlowitz HC. Sylvicultura oeconomica. Anweisung zur wilden Baum-Zucht. Freiberg: TU Bergakademie Freiberg und Akademische Buchhandlung, 1713. Reprint edited by Irmer K and Kießling A, 2000.

[8] UN World Commission on Environment and Development. Our Common Future (Brundtland Report). Oxford: Oxford University Press, 1987.

[9] Maine T. Towards a Metric of Sustainability, in: B. Douglas, B. Furnass and J. Goldie, (Eds.), ISOS online conference - In Search Of Sustainability, 2003.

[10] Dalal-Clayton B. Modified EIA And Indicators Of Sustainability: First Steps Towards Sustainability Analysis. Environmental Planning Issues 1993: 1: International Institute For Environment And Development.

[11] Odum HT. Environmental Accounting: Emergy and Environmental Decision Making. New York: Wiley, 1996.

[12] Campbell DE, Brandt-Williams SL, Meisch MEA. Environmental Accounting Using Emergy: Evaluation of the State of West Virginia. United States Environmental Protection Agency. 2005.

[13] Jorgensen SE, Nielsen SN, Mejer H. Emergy, environ, exergy and ecological modelling. Ecological Modelling 1995: 77: 99-109.

[14] Brown MT, Ulgiati S. Emergy Evaluation of the Biosphere and Natural Capital. Ambio 1999: 28 (6).

[15] Zhao S, Li Z, Li W. A modified method of ecological footprint calculation and its application. Ecological Modelling 2005: 185 (1): 65-75.

[16] Coiera E. Putting the technical back into socio-technical systems research. Int J Med Inform 2006: to appear.

[17] Moore J. A new scheme for data sharing - Semantic interoperability aims to ease data sharing among disparate health systems. Government Health IT 2006: 13. Nov. 2006.

[18] Chen R, Klein G. The openEHR Java Reference Implementation Project, accepted for Medinfo 2007.

[19] Iversen KR, Grotan TO. Socio-technical aspects of the use of health related personal information for management and research. Int J Biomed Comput 1996: 43 (1-2): 83-91.

[20] Shabo A. A global socio-economic-medico-legal model for the sustainability of longitudinal electronic health records - Part 2. Methods Inf Med 2006: 45 (5): 498-505.

Address for Correspondence

Dr. Sebastian Garde
Health Informatics Research Group, Central Queensland University, Austin Health, Melbourne, VIC, Australia
Phone: +61 (0)3 9496 4040
s.garde@cqu.edu.au
http://healthinformatics.cqu.edu.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Bermuda Triangle or Three to Tango: Generation Y, e-Health and Knowledge Management

Kwang Chien, Yee

Research Fellow, eHealth Services & research Group, School of Information Systems, University of Tasmania, Australia.

Abstract

Generation Y workers are slowly gathering critical mass in the healthcare sector. The sustainability of future healthcare is highly dependent on this group of workers. This generation of workers loves technology and thrives in stimulating environments. They have great thirst for life-experience and therefore they move from one working environment to the other. The healthcare system has a hierarchical operational, information and knowledge structure, which unfortunately might not be the ideal ground to integrate with generation Y. The challenges ahead present a fantastic opportunity for electronic health implementation and knowledge management to flourish. Generation Y workers, however, have very different expectation of technology utilisation, technology design and knowledge presentation. This paper will argue that a clear understanding of this group of workers is essential for researchers in health informatics and knowledge management in order to provide socio-technical integrated solution for this group of future workers. The sustainability of a quality healthcare system will depend upon the integration of generation Y, health informatics and knowledge management strategies in a re-invented healthcare system.

Keywords:

Generation Y, health informatics, knowledge management, human resource development, technology

Introduction

The healthcare system is at a cross-road. In the next few decades, this industry is undergoing the most extensive transformation ever seen in any industry in history [1]. There are many internal and external factors that drive changes to the healthcare system; some of these factors are very relevant to health informatics and knowledge management:

1. The population that the current healthcare structure serves is rapidly ageing [2]. Patients are getting older and they are more likely to have multiple medical problems with increasing number of prescribed medications. A recent study from Australia has shown that the average age of patients admitted to medical wards is 74 and the average number of prescribed medication is 9.6 [3]. The problem of ageing, increasing complexity and polypharmacy is only going to worsen with time.
2. Secondly, the healthcare system is experiencing rapid subspecialisation, with increasing utilisation of technology. This creates the need for increase communications and information exchanges among various healthcare professionals.
3. Thirdly, advances in technology have become so rapid that it is difficult for healthcare professionals to keep up to date with all areas [4]. Therefore, point-of-care decision support and evidence based guidelines will become increasingly important.
4. The proliferation of knowledge management models and health informatics innovation aims to provide the right information at the right place at the right time. The delivery of information to clinicians at the bedside, however, is often not filtered and therefore an increasing amount of information reaches the desks of clinicians everyday.
5. We have an ageing workforce that needs replacing. This ageing workforce is now being slowly replaced with generation Y workers. Generation Y workers have very different expectations and work practices that a generational gap is widening within the healthcare workforce.

This paper examines the characteristics of generation Y workers applicable to the healthcare system, drawing on the experience from other industries. It then presents a discussion about the potential benefits and pitfalls of working with generation Y in the areas of health informatics and knowledge management. The paper argues a likely adverse outcome if the healthcare re-invention process fails to take these social-cultural issues into account. Finally, the paper proposes some simple rules to deal with generation Y in the health informatics and knowledge management areas and argues that a successful re-invention with generation Y will produce a sustainable healthcare system.

Generation Y and healthcare workplace

Generation Y is commonly defined as those born after 1978 [5], although the exact cut off year is arguable [6]. This generation of young workers has very different understanding and expectation of the world. They are creating a big impact in other industries [7]. As they are now getting

into their 20's, their impact on the healthcare system is slowly being acknowledged. Many characteristics of generation Y have been described and reviewed elsewhere [5-6]. This section will focus on the characteristics which are relevant to electronic healthcare implementations and knowledge management

Technology savvy

Generation Y grows up, surrounded by technology [5]. They see the diffusion of computer from academic research centres into everyday life. They experience the impact of technology not only on their activities of daily living but also career creation and financial stability. Technology has not only transformed their lives, but also provided limitless opportunities for generation Y from "Yahoo" to "Youtube". The healthcare system is unfortunately an industry that is slow to take up technology. When it does utilise technology, however, the story of failure is common. In fact, 75% of big IT projects in healthcare fail [8].

Stimulation and challenging tasks

Generation Y hates routine tasks [9]. They want to have fun during their routine work. They thrive in stimulating and challenging conditions [5]. The industry revolution and now the information age revolution have seen the automation of most manufacturing jobs. Unfortunately, most of the routine day-to-day jobs in healthcare system is still carried out by human beings. Generation Y loves multi-tasking, especially using multiple technological devices. This is a challenging issue in healthcare, especially when quality and safety is taken into consideration.

Information gathering and presentation

Generation Y gather information with lightning speed. They are street smart and they gather information by the fastest means, including the utilisation of technology, but most likely, they obtain essential information through social networking and mobile messaging from their friends [5-6]. Generation Y wants simple information presented in simple language. In generation Y terms: "U get info U wnt 2 r mates only. :)". Unfortunately, this culture does not fit into the formality and hierarchy of healthcare system [4] and generation Yers are perceived to often bypass the usual protocols.

Education and training

Education and training is part of generation Y's life [5]. They are the most educated generation of all time. They, however, learn in a different way. Generation Yers want on-demand, ubiquitous and relevant education and training. These education and training needs are often supported by technology. The healthcare system is however, based heavily on an apprenticeship model [10]. Education and training provided are often irrelevant to their perceived immediate needs.

Fluidity of workforce

Generation Y workers like variety. They are likely to change jobs or institutions frequently [5]. The work model of students working their way through to consultants at the same hospital is being challenged. Generation Yers feel bored when they need to remain in one job for a long time. This creates significant problems in the medical workplace as every hospital is slightly different. While explicit knowledge representation has been emphasised, rising on the backdrop of evidence-based medicine, representation and delivery of tacit knowledge which until now has largely been ignored, becomes difficult because of the fluidity of workforce.

Potentials for health informatics and knowledge management with generation Y

There is significant potential for health informatics and knowledge management to flourish with the increasing critical mass of generation Y in healthcare workforce. Most generation Y healthcare professionals are computer literate and love technology. This section will explore the opportunity for health informatics and knowledge management, working with generation Yers.

Health informatics

Generation Yers swim in technology since birth. They believe in technology! Generation Y finds it easy to adapt to changes and utilise new technology. Given the ageing population, the increased complexity of medical disease management and increased number of prescription medications, implementation of an electronic health record is the only sustainable future. The implementation of an electronic healthcare system will not require significant upskilling of generation Y. In effect, generation Y will probably suggest the use of technology for most of their routine work. They want technology to fit into their needs and their world. This creates more needs for information systems researchers, not only to search for techno-social integrated solution, but also to provide a direct bridge between technology evolution and medical workforce evolution.

The speed of technology roll out offers the necessary stimulation for generation Y to remain interested in the healthcare system and to participate in the redesigning process of the healthcare system. While this might sound like a daunting task to many, generation Yers will find it challenging and they will want to be part of this evolution.

Their familiarity of search engines, multi-tasking and electronic storage system from young age means that generation Y will adapt to electronic information much easier. They will be able to find information quickly through the maze of digital coding. As generation Yers like to improve the efficiency of performing common tasks, they will provide useful suggestions for future healthcare revolution. Their suggestions will likely to be not only useful, but also practical!

Knowledge management

The traditional knowledge management model discusses the stages of elicitation, representation, sharing, evolution and delivery of knowledge [11]. While the theory of knowledge management seems to suggest a distinction between tacit and explicit knowledge [12], in the medical field, tacit and explicit knowledge seem to be inter-con-

vertable, especially with clinical disease management [11].

The request for on-demand knowledge representation by generation Y and the increasing fluidity of the workforce with generation Y are likely to rapidly increase the need for better knowledge management within the healthcare system. The distinction between tacit and explicit knowledge management will become more evident with generation Y.

Firstly, generation Y wants rapid and on-demand access to education, training and clinical decision support systems. The proliferation of electronic resources makes the integration and development of up-to-date guidelines difficult. The rapid expansion of subspecialisation and research means that guidelines are almost out of date the moment consensus has been reached. There needs to be a better way to elicit knowledge electronically. Third generation search machines and artificial intelligence will be essential if we are to gain the confidence of generation Y.

Secondly, generation Yers will not be interested in learning knowledge irrelevant to their current needs. Therefore, they rely heavily on technology to deliver up-to-date, on-demand knowledge representation. As knowledge representation and sharing become more rapid and cross-institutional, the difficulties in version management will increase. The problem not only relates to updating the current knowledge representation, but also ensuring the removal of older versions from searchable knowledge data-base [11].

Thirdly, the delivery of knowledge and information retrieval systems for generation Y will have to be intuitive. Ubiquitous feedback algorithms that provide "related" information for generation Y will need to be relevant, individualised and user-centered.

Finally, the distinction between explicit and tacit knowledge will become obvious as generation Y moves from one institution to the other. The medical workplace thrives on individuality. There are not only issues with different guidelines for clinical disease management, more importantly, significant differences exist in leadership, management skills, radiology ordering, pathology ordering, pharmacy dispensing as well as interpersonal communication. The tacit knowledge which enables this has previously been transferred through the process of socialization. With the increasing number of short-term employment, part-time and locum workforce [13], the representation of tacit knowledge will have to be delivered to the workforce rapidly, allowing another dimension of knowledge management to flourish.

Pitfalls for health informatics and knowledge management with generation Y

While there are potentials for health informatics and knowledge management to expand, there are also pitfalls working with generation Y.

Problems with technology

Generation Y has very high expectations of what technology should deliver. While the technology might be available, there are multiple issues to be considered, such as ethics, security, data integrity, cultural factors, environmental factors and social factors [14]. It will be difficult to communicate the relevance of these factors to generation Y.

The healthcare budget is limited. It is not possible to fund cutting edge technology all the time. Low end technology will probably be as good and as efficient as cutting edge technology in performing some of the routine clinical jobs. While from a health economics point of view, it is imperative that these are taken into account, it will not fit into generation Y's culture of "being cool!"

Generation Y is already creating a generation gap in leadership, clinical governance and work-life balance [15]. Their familiarity with technology and their desire to utilise technology, if not managed appropriately, will lead to widening generation gap and disharmony in the healthcare workplace.

The implementation of new technology in a complex system often leads to multiple unforeseeable problems. The healthcare system is arguably the most complex of all systems. While there are socio-technical integrated solutions, there are still more failures that successes [14]. While the experience of generation Y in technology evolution might assist them in their adaptation to technology, it might also foster the misunderstanding that technology will fix all problems. The complexity of the healthcare system and therefore the difficulties in Information Technology (IT) implementation might not be easily comprehensible to generation Y.

Generation Y might want to be involved and lead in early stages of IT implementation. They, however, lack the necessary experience of project management. Their nature of experimentation and overestimation of their capacity are significant problems with IT implementation in healthcare as safety is often not the priority for generation Y.

Problems with knowledge management

While generation Yers are very good at information gathering, their information presentation is difficult to be shared with outsiders. The SMS messaging forms a unique language structure. This language is spreading rapidly, including its utilisation in healthcare messages for generation Y [16]. While the SMS language is gaining acceptance, including formal examination in New Zealand, many healthcare workers might struggle to understand the language. Should we use SMS language to transmit medical data? This is a great dilemma that we will need to resolve soon.

While it is often possible to retrieve on-demand information and to receive on-demand education, one needs to have a certain level of knowledge and competency in the medical world. Furthermore, the hierarchical structure assumes certain functions for each individual within that organisation. The on-the-spot clinical decision support and

knowledge representation for generation Y will challenge the basic fundamental assumption of the medical world. This has the potential to create significant conflict within the medical profession.

While knowledge elicitation and representation might be achieved regarding disease treatment, the tacit knowledge of "know-how" might realistically be very difficult to be elicited and represented. This is especially so for interpersonal interaction, leaderships and management styles. It might cause disastrous consequences if attempts were made to elicit, present and share this sensitive tacit knowledge.

Three to tango: Integration

Is the combination of generation Y, health informatics and knowledge management a Bermuda Triangle that should never be entered? Is it possible to integrate generation Y, health informatics and knowledge management together to form an irresistible force to transform the healthcare system? This paper argues that it is possible to integrate the three together to form a tangible connection to build a sustainable future healthcare system. This section will provide health informatics and knowledge management researchers and practitioners some simple guidance to work with generation Y.

A recent paper recommends ten commandments for IT designs in healthcare[17]. These ten commandments are:

1. Speed is everything;
2. Anticipate needs and deliver in real time;
3. Fit into the user's workflow;
4. Little things can make a big difference;
5. Recognise that physicians will strongly resist stopping;
6. Changing direction is easier than stopping;
7. Simple interventions work best;
8. Ask for additional information only when needed;
9. Monitor impact, get feedback and respond;
10. Manage and maintain your knowledge based systems.

These ten commandments are proposed in order to achieve socio-technical integration for IT implementation, especially decision support system in the current healthcare system. The future, however belongs to generation Y! We therefore need to take into account the characteristics of generation Y when building the future healthcare system. This paper suggests 10 different commandments for IT implementation in the healthcare system with generation Y. This is based on the synthesis of the understanding of generation Y in the literature as well as the author's experience of working with generation Y.

1. **Everything is speed**

   Generation Y wants outcomes and wants them fast. Therefore, from design and communication to feedback and improvement, speed is the key. Otherwise, generation Y will get bored and move on.

2. **Find out their needs and deliver wirelessly.**

   It is very important to find out the needs of generation Y and to involve them early in the design phase. While some of their suggestions might be difficult to achieve, their early involvement motivates them to help you.

3. **Empower juniors to change workflow through health informatics.**

   While IT implementation that fits into the workflow is desirable, IT implementation that drives the change of workflow is the future! It is a golden opportunity to empower generation Y to suggest current work flow changes, and deliver these changes through IT implementation.

4. **What are little things and big things?**

   There is nothing that is big or little in the minds of generation Y, as they think globally but act locally. Therefore, it is essential to ensure that all the important issues are addressed despite the perceived triviality by the IT team.

5. **Recognise generation Y wants progress all the time!**

   Generation Yers get bored easily. It is therefore of upmost importance to have a clear version management strategy and communicate the plan to generation Y. If improvements are made based on feedback from generation Y, communicate that clearly.

6. **Don't stop them doing the wrong thing; ask them to do the right thing more often.**

   The implementation of IT project, especially clinical decision support and guidelines will inevitably change the underlying process. The challenge to do the right thing more often is much more powerful than a condemnation to stop doing the wrong thing.

7. **Simple intervention that multi-tasks, works best**

   Generation Y loves multi-tasking and they are good at it. While intervention needs to be simple, it also needs to provide generation Y with the potential to multitask.

8. **Don't ask for information, "google" for information.**

   Generation Y is good at searching for information, especially electronic information. It is essential that IT programs are intuitive and manage to "google" information that is already available in database rather than asking generation Y repeatedly for the same information.

9. **Check, evaluate, change or move on!**

   Generation Y learns from experience and errors and they expect others to do the same. It is therefore important to share the experience and errors of IT implementation with them. In the event of unexpected outcomes, evaluation and changes need to be acted upon quickly. In the event that improvement is impossible, a direct response to generation Y is essential and then move on!

10. **On-demand knowledge representation in their language.**

Knowledge representation, either tacit or explicit knowledge should be ubiquitous and on-demand. While the SMS language might not be the official language and might be difficult to understand, it is the language that will communicate the best with generation Y!

Conclusion

In the next few decades, the healthcare system is facing the most massive transformation of any industry. Among all the challenges, the replacement of the ageing workforce by generation Y workers will create a massive impact. Generation Y workers have very different characteristics. Their inclination to utilise technology and to create life experience generates an opportunity for health informatics and knowledge management to flourish. There are, however, pitfalls working with generation Y. This paper proposed ten recommendations to integrate generation Y with health informatics and knowledge management for a sustainable future healthcare system:

1. Everything is speed;
2. Find out their needs and deliver wirelessly;
3. Empower juniors to change workflow through health informatics;
4. What are little things and big things?
5. Recognise generation Y wants progress all the time!
6. Don't stop them doing the wrong thing; ask them to do the right thing more often;
7. Simple intervention that multi-tasks, works best;
8. Don't ask for information, "google" for information;
9. Check, evaluate, change or move on!
10. On-demand knowledge representation in their language.

Acknowledgments

Thanks to all generation Yers who worked with me, inform me, educate me and enlighten me on their journey towards excellence. "U R the BEST!"

References

[1] Fitz-Enz J. Industry change = internal change. HRD/OD, Am. Management Association International. 1998: 1: 5.
[2] Lozada M. The health occupation boom. Vocational Education Journal 1995: 70: 34-7.
[3] Yee KC, Barbour L. Improving medication safety at the interface: The role of pharmacist versus information technology in the real world. Program and abstracts of Improving Medication Safety Conference: Sharing the lessons learned , Sydney, New South Wales, Australia. 2006.
[4] Benson S and Dundis SP. Understanding and motivating health care employees: integrating Maslow's hierarchy of needs, training and technology. J Nurs Manag 2003: 11: 315-20.
[5] Sheahan P. Generation Y: Thriving and Surviving with Generation Y at Work. Victoria: Hardie Grant Books, 2005.
[6] Huntley R. The World According to Y: inside the new adult generation. New South Wales: Allen & Unwin, 2006.
[7] Salt B. Y the younger generation is so different. The Australian. 16th November 2006. p. 26.
[8] Littlejohns P, Wyatt JN and Garvican L. Evaluating computerised health information systems: hard lessons still to be learnt. BMJ 2003: 326: 860-3.
[9] Workout what Gen Y wants. The Australian, Weekend Career One. 25th-26th November 2006. p. 1.
[10] Gorman D, Sir J Scott. Time for a medical educational change in time. Intern Med J 2006: 36: 687-9.
[11] Peleg M and Tu S. Decision support, knowledge representation and management in medicine. Methods Inf Med 2006: 45 (s1): S72-S80.
[12] Nonaka I, Takeuchi H. The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation. London: Oxford Unviersity Press, 1995.
[13] Skinner CA, Riordan RL, Fraser KL, Buchanan JD and Goulston KJ. The challenge of locum working arrangements in New South Wales public hospitals. Med J Aust 2006: 185: 276-278.
[14] Wears RL and Berg M. Computer technology and clinical work: Still waiting for godot. JAMA 2005: 293: 1261-3.
[15] Yee KC, Tsunami to SurfWave: Toward harmony in generation Y gap challenge. Program and abstracts of Towards Harmony in the Health Workplace Conference: Practical approaches to managing difficult interactions and preventing the escalation of conflict and grievances , Sydney, New South Wales Australia. 2006
[16] Bisset K. Get set for 24/7 help with sextxt. The Australian, Weekend Career One 2nd-3rd December 2006. pp. 19.
[17] Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, Spurr C, Khorasani R, Tanasijevic M and Middleton B. Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. J Am Med Inform Assoc 2003: 10: 523-30.

Address for correspondence

Dr. Kwang Chien Yee
eHealth Services & Research Group
School of Information Systems
University of Tasmania
Hobart, Tasmania, Australia.
Email: kwang.yee@dhhs.tas.gov.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Health Service Organisations and Professionals: An Information Systems Model for Transforming the Nexus between Accreditation and Practice

Jo-Anne Kelder

School of Information Systems, University of Tasmania, Australia

Abstract

This paper presents a qualitative research approach used to generate data and theoretical insights for information systems design in a highly regulated health service organisation. Ethnographical and sociological analytic techniques were used for sensemaking in the domain, and then identifying key structures affecting the conduct of the health service. A key aspect was taking a futures perspective of possibilities for evolution of these structures. From this, a continuum model representing the nexus between accreditation and practice was developed, varying the degree of integration of information systems for organisation- and individual-level accreditation. The paper uses the model to discuss the implications of possibilities for design in organisations where accreditation requirements have implications for work practice and information systems design.

Keywords:

Accreditation, health informatics, organizations.

Introduction

National health care systems are currently subject to transformational change drivers that threaten sustainability; health informatics is preoccupied with the question of how to model and design sustainable technical systems in complex, dynamic socio-technical health settings. This research assumes that a 'whole of system' design approach is needed to avoid failure in health service ICT design and implementation [1].

Health service organisations and health professionals who work for them are increasingly subject to quality assurance processes and accreditation requirements. But recent research raises questions on whether the abundance of very costly accreditation processes delivers better health systems [2].

Quality frameworks for organisations consist of standards for operation and metrics for measuring performance as the basis for determining accreditation status. ICTs have an integral role in quality assurance of health services: they enable collecting and managing vast quantities of data onto digital information systems and interrogation of the system to evaluate an organisation's compliance to the standards [3].

This paper presents the outcome of a research approach used to generate data and theoretical insights for information systems design. The research setting was a highly regulated health service organisation, BreastScreen Tasmania (BST). BST is a member of BreastScreen Australia (BSA), the accreditation organisation for the Australian breast screening Program. Each member organisation is accredited on the basis of compliance to 176 standards, comprising the 'best practice' according to experts from professions including radiography, radiology, pathology and surgery as well as best practice in data management, management of client psycho-social needs and other contributing professions.

The research method has three phases and draws on established ethnographic and sociological techniques and frameworks for generating and structuring data from the setting. The criterion for selection is their capacity to accommodate and deliver insight into the interactions and relations between the people, the place and the things (PPT) in the setting. Many such PPT frameworks have been fruitfully used in medical/health informatics research projects [4, 5].

PPT frameworks used for applying initial structure to the data were Distributed Cognition theory (Dcog) [6] which focuses on people-artefact interactions developing over time as a cognitive system, and Communities of Practice theory (COP) [7] which directs attention to the existence of mechanisms for sharing and distributing knowledge and information (people-brokers, artefacts-boundary objects). Activity Theory (AT) [8] was applied for its capacity to engender thinking about PPT interactions as affected by tools, rules and division of labour for a community activity. It facilitates representing multiple perspectives (different subjects with different objectives) and helps identify misalignment between the possible objectives of an activity and elements used in the conduct of the activity

Method

This method was designed for research in complex socio-technical settings. It varied the use of ethnographic and sociological analytical techniques over three phases of data collection and analysis. The first two phases of the method can be applied to any complex socio-technical domain for the purposes of making sense of the domain and creating structures understanding it. The outcome of third phase is specific to the research setting, but the principles of the method can be applied more generally.

Phase one: Sense-making

The first phase was a sense-making exercise for the researcher. The initial research questions were: 1) *How to approach IS research in complex and sensitive (highly social) domains?* and 2) *What are the relationships between the situation of the work place, digital information technology/ other artefacts in the domain and the ways people perceive and do their work?*

This phase was deliberately exploratory and involved several months of immersion in the research setting. Ethnographical techniques were used (field observations, semi-structured interviews, document collection and iterative analysis) to explore the setting; PPT techniques (from Dcog, COP and Activity Theory) were used to initiate understanding the setting by applying some structure to the raw data, and directing the researcher's attention to specific types of inter-relationships between the people, place and things in the domain.

Immersion involved observing people at work and conducting semi-structured interviews for insights into what was observed. Informal and observation-associated comments and conversations with staff members were recorded as field notes. The researcher reviewed over 200 organisation documents and scanned the wider environment using automatic alerts of electronic content relevant to the health service context. BST community education and cancer policy development staff members also made available their information resources.

A technique used for exploring the setting was to follow multiple client trajectories through the series of interactions with the organisation to identify *client* perspectives and assumptions about why and how they engage with the health service [9]. Field notes were taken for 24 clients having a routine mammogram, and seven clients attending an assessment clinic.

Client trajectories for personal interactions with BST were then extended by tracing the trajectory of artefacts and people connected with client interactions beyond the immediate context. For example, the client record (CR) was moved by trolley between the clinic area and the data management area. Clinic data entered onto the physical CR had to be entered onto the Client Information System (CIS). The trajectory of CR-CIS interactions gave a rich data set of coordination and breakdowns of activities, problem situations and multiple interpretations involving individuals, work teams and the organisation.

PPT approaches share techniques from ethnography (field notes, interviews, document collection) but collect data from their own perspective. Collecting data without particular reference to a PPT theory, and experimentally applying the differing theoretical frameworks enabled exploring a variety of possible structures for framing understanding the data.

Phase two: Identifying problems

This phase of the research focused on identifying problem themes. Theme discovery was facilitated by paying attention to aspects of the setting connected to problems in conducting the organisation enterprise (screening women for breast cancer). Themes were identified by iteratively applying PPT techniques to gain theoretical insight into the setting and by using the trajectories technique to uncover perspectives and structural relations not identified by specific PPT frameworks.

A theme was selected for further investigation according to its capacity to express a significant problem-situation, affecting at a structural level the information systems and work practice systems design in the research setting. It was the basis for new research questions in Phase 3. For BST, the theme was *the relation between accreditation and practice.* And the research questions: 3) *What are the consequences of the way accreditation is framed for people and the artefacts they use in their work situation?* and 4) *What relationships are included in the nexus between accreditation and practice and what are the implications for IS design?*

Phase phree: Researching an emergent theme at the level of structure

In phase three, the theme was researched iterating a sequence of techniques designed to conceptualise and model the structures of social organisation expressed in the theme. Ethnographic observations were extended by following trajectories of interactions beyond the immediate observation context followed by conceptual modeling of the data at the level of structure and testing the structural models via further ethnographic observations.

Ethnographic observations focussed on identifying the elements and relations in PPT interactions connected to the problem-theme; trajectories of interactions beyond the organisation boundary were followed to identify the contexts and drivers for different perspectives found within the organisation.

Conceptual modelling of the data was used to identify key structures shaping and constraining the attributes of the organisation. Some of these structures could be derived from the PPT theories used in the previous phases; some were created using data that did not fit the theories. COP theory and AT provided constructs for identifying key elements in the data and modelling their structural relations. Trajectories pointed to data sets and relations that required creating additional constructs to theorise them adequately.

This sequence of observations modelling testing the model was iterated to produce models of the elements and relations structuring the relationship between being accrediting for breast screening practice and how the enterprise was conducted within BST.

The nexus between accreditation and practice was conceptualised in four related models: 1) organisation-level accreditation; 2) individual-level accreditation; 3) the nexus between accreditation and practice for organisation-level and individual-level accreditation, mediated by boundary maintenance (Figure 1) and 4) a context diagram setting out the artefact and membership (information) connections between individuals, networks of practice who accredit them, an organisation conducting a health enter-

prise and an organisation whose role is to accredit that organisation (Figure 2).

Application of the models - Futures perspective

The third and fourth models were used as the basis for imagining alternative structures for the future, occurring either by natural evolution or by design. The technique of trajectories was used to investigate the changes over time of visions, negotiations, decisions and activities that led to the current partially integrated structures for legitimating individuals and organisations for conduct in health-related activities. An integrated structure for accreditation was based on current change drivers and desirable attributes for a future health system .

A continuum model of the *accreditation – practice nexus* (Figure 3), varied according to the degree of integration of individual- and organisation-level accreditation systems provided the focus for envisaging structures for accreditation. The properties of the IS and work practice systems were derived for the integrated end of the continuum by considering possibilities for data sharing between the different entities constituting an.

Results

This section outlines the constructs and the structural models developed in Phase 3.

Constructs

COP analysis provided models of systems of work practice within the clinic at individual-level that highlighted the role of *membership* in a competent *community of practice* for establishing competence at the level of work practice. Trajectories of professions within the multi-disciplinary team at BST established that health professionals are accredited in a social system of *multi-membership*: entry into a specific work domain (and COP) is dependent on prior accreditation as a continuing member of a professional association, a *network of practice* (NOP) [10], with membership obligations such as demonstrating continuing professional development (CPD). PPT interactions observed around the nexus of accreditation and practice were 'collegial' that is, operating at a local level and based on personal relations and trust; artefact use was limited; individuals were responsible for collecting and maintaining the evidence of *competency* for membership in the NOP, and for their reputation in the COP.

AT was applied to the data at the level of activity because it addresses 'rules' and 'tools' operating in a given activity and the research had documented organisation effort in ensuring it complied with the national accreditation standards (NAS). The data also evidenced influence of the accreditation requirements originating from BreastScreen Australia (BSA) on practice within the breast screening clinic context. The *accrediting organisation* designed artefacts to monitor and evaluate the competence of BST at organisation-level and required that *accreditation artefacts are naturalised into the practice* of the breast screening COP. All policy and procedure manuals and the client record in particular, were designed and written to comply with the NAS for organisations accredited for membership in the national Program.

New constructs were created to describe and model the organisation-level response: a set of activities designated *boundary maintenance*, to enable coordination and compliance from the breast screening *enterprise organisation* (BST) to the *accrediting organisation* (BSA). The concept of *trajectory into a community of practice* was also constructed to capture the two aspects of membership linked to individual membership of a *network of practice*. That is: entry into the breast screening COP is predicated on NOP membership, but membership of a COP is negotiated and established over time as an individual demonstrates competence in applying NOP skills and knowledge in the joint conduct of the enterprise. This is a social construction of accreditation.

Conceptual models

Figure 1 represents the nexus between accreditation and practice in situations of different systems for organisation-level and individual-level accreditation, mediated by boundary maintenance. The boundary maintenance required to bridge the disjunctions between the individual-level social construction of accreditation (mediated by an individual's membership of both the community of professionals with whom they work and the professional community of their vocation) and the organisation-level artefact construction of accreditation (mediated by information contained on client records (CR) converted to information on digital client information systems (CIS)).

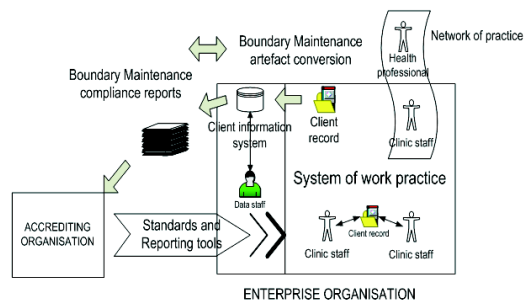


Figure 1 - Accreditation mediated by boundary maintenance

Boundary maintenance is the activity set required within the organisation to convert the information artefacts used for practice (the CR for each woman registered with the breast screening program) into information artefacts for organisation accreditation (the CIS which aggregates the data of all clients into statistics such as screening rate and cancer detection rate for the target population which are used to generate reports used to determine BST's compliance to the NAS).

Accreditation of an organisation is conducted largely via digital information artefacts processing data and generating reports. Membership in the BSA Program is contingent on successful implementation of the NAS into all aspects

of an organisation's practice. This naturalisation of standards is supported by naturalising BSA's artefacts for use in the organisations engaged in the breast screening enterprise. This takes the form of templates, the standards and accompanying manuals; it includes the effect of requiring that work artefacts developed within the organisation (CR, policy and procedure manuals) comply with the NAS. It is an artefact construction of accreditation.

Accreditation for individual health professionals is separated from organisation accreditation systems. Individuals maintain membership with their professional organisation by participating in approved continuing professional development activities and being able to provide documentary evidence.

Figure 2 represents the connections (evidenced by information transfer processes) between four major s in the social organisation of accreditation in the health context: individuals, networks of practice, health enterprise organisations and the accrediting organisation.

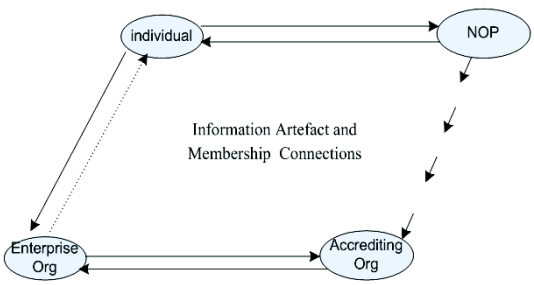


Figure 2 - Accreditation infrastructure entities

The accrediting organisation, BSA, authorises the National Accreditation Standards as current best practice for the professions engaged in its Program. Networks of Practice (NOP) connect with the accrediting organisation by their representatives on committees determining standards, and when individuals are invited to participate in a site accreditation visit to audit the data and practice of an enterprise organisation. BSA and NOP information artefacts and systems are separate.

The standards and measures expressed in the NAS act recursively to educate and support their adoption as standards of best practice generally within the NOP and across the health domain beyond the immediate context of breast screening.

BSA primarily interacts with member breast screening organisations via the output of digital artefacts. Individual professionals working for BST create data which is aggregated for demonstrating the organisation is a competent member of the Program. But they have limited access to data relating to their individual work practice. Radiologists are given feedback on key indicators such as the cancer detection rate and discuss this data privately or in the context of multi-disciplinary meetings with surgeons and pathologists to improve the knowledge and skills of everyone involved. Other members of the clinic community of practice rely on information generated and maintained within their local system of work practice, including feedback from colleagues, comments by clients and the data on the client record.

Future thinking for IS design - A continuum model

Figure 3 represents a continuous spectrum of possible accreditation structures for establishing the legitimacy of health service organisations and individual professionals working within them: each possible structure associated with a system of work practice and information systems. Work practice systems (WPS) and information systems (IS) for any point on this continuum will have certain properties which can be identified by systems analysis techniques such as business process analysis or PPT analyses. Transformation of the structures and associated systems requires taking into account those properties and identifying the issues and drivers for transition from one nexus to another.

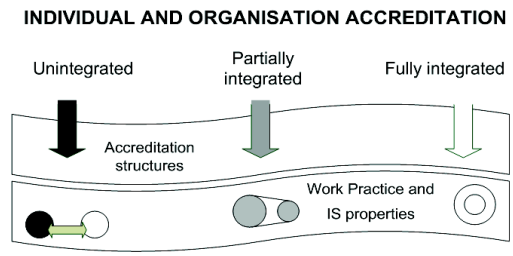


Figure 3 - Continuum model accreditation-practice nexus

Unintegrated accreditation structures are characterised by separate artefacts and information systems: the information system for collecting data for organisation level accreditation is not used by individuals in their work or for their accreditation. Separate WPS and IS are used in the conduct of the enterprise and individual accreditation activities are also separate from their work practice. Data from practice artefacts (e.g. the CR) must be added onto accreditation artefacts (e.g. the CIS) by additional processes. This boundary maintenance activity requires additional staff members and organisation resources and is a locus of breakdowns in organisation activities.

Partial integration occurs via the naturalisation of accrediting organisation standards and other artefacts into the enterprise organisation's WPS. Competent use of these artefacts and compliance with standards becomes part of the individual's identification of what it means to be an accredited professional in that environment. Partial integration at BST is evidenced by use of data on cancer detection rates that is measured for organisation-level accreditation and also provides feedback for individual radiologists. However, boundary maintenance activity is still required to utilise data for all levels of accreditation. For example, a BSA standard for accreditation data at organisation-level includes evidence of individuals attending multi-disciplinary meetings, data which can also be used as CPD evidence for continuing eligibility for NOP membership. Here, the systems are separate and require human work to collect and format into reports.

Where accreditation systems are fully integrated, a single artefact acts as a boundary object [11] for all entities in the accreditation infrastructure (Figure 2), providing information sufficient and appropriate for the needs of each entity. If the IS used within the organisation in the course of conducting the enterprise also takes the work practice data to measure competency for individuals and the organisation, *and* that IS artefact and its measurement is accepted by the NOP, this displaces the individual's work of demonstrating meeting NOP membership requirements onto the IS that records information about their work. Boundary maintenance activity is eliminated.

Discussion

Unintegrated, or partially integrated systems for organisation and individual level accreditation are problematic for health services such as BST at the level of WPS and IS design. Boundary maintenance (BM) activity at the intersection between practice artefact (CR) and accreditation artefact (CIS) constantly breaks down, with implications for the organisation's accreditation and resource requirements. Figure 2 can be used to consider alternative information flows and artefacts to support accreditation and practice.

Designing a single artefact that meets both practice and organisation accreditation requirements removes the need for resources for boundary maintenance. This is predicated on the organisation standards being incorporated into the relevant standards and requirements for professional bodies, thus removing the individual's burden of fulfilling distinct sets of requirements for multiple memberships. It is also predicated on appropriate permissions for access to data being agreed.

Australia's demographic profile points towards a continuing shortage of qualified health professionals paralleled by an increasing client base. Digital mammography is quicker than film and provides data that can be linked to an electronic client record and performance record of the radiographer and radiologist. Providing feedback to practitioners on individual skills and knowledge competence enables communication for improvement. This happens successfully with radiologists in the national Program. Information of individual radiologists' performance in their work (which contributes to the organisation accreditation) is provided and used as an opportunity for the team to discuss and take steps to improve knowledge and skills. Individual performance linked to organisation-level competence status has contributed to a COP-based drive for excellence and improving the quality of client care.

Integrated accreditation systems which provide real-time performance feedback to professionals and their NOP may provide a less time-consuming and costly process of accrediting professionals for work – relevant to the debate of importing health professionals trained overseas to work in Australia, or of exporting digital images to radiologists overseas and wishing to be assured of the quality of their work.

Conclusion

This method produced a set of conceptual models for thinking about the nexus between accreditation and practice in the health care setting. It enables consideration of the impact of different structures for accreditation on the design of information systems and work practice systems for a given health service enterprise, and a vision for integration that can contribute to the transformation of the wider national health system. In particular, identifying the location of an organisation on the continuum model is critical for understanding the constraints on ICT design for a given nexus: design in the context of unintegrated accreditation systems will be constrained to provide for boundary maintenance between the two.

Acknowledgments

The researcher thanks her supervisors, A/Prof. Paul Turner and Dr John O'Neill, the SIT CRC for their support. She especially thanks the staff members at BreastScreen Tasmania. This paper presents her doctoral research and is her own work.

References

[1] anon. Key Issues and Future Directions in e-Health-perspectives on 2010 and beyond: User Environment Program, School of Information Systems, University of Tasmania; 2004.

[2] Braithwaite J, Westbrook J, Greenfield D, Pawsey M, Naylor J, Iedema R, et al. A prospective, multi-method, multi-disciplinary, multi-level, collaborative, social-organisational design for researching health sector accreditation. BMC Health Services Research 2006;6:113-123.

[3] Clarke AE. Situational Analyses: Grounded Theory Mapping After the Postmodern Turn. Symbolic Interaction 2003;26(4):553-76.

[4] Nemeth C, O'Connor M, Klock PA, Cook R. Discovering Healthcare Cognition: The Use of Cognitive Artifacts to Reveal Cognitive Work. Organization Studies 2006;27(7):1011-1035.

[5] Bate SP, Robert G. Knowledge management and communities of practice in the private sector: Lessons for modernizing the national health service in England and Wales. Public Administration 2002;80(4):643-663.

[6] Hutchins E. Cognition in the Wild. Cambridge, Mass: The MIT Press; 1995.

[7] Wenger E. Communities of Practice: Learning, Meaning and Identity. Cambridge: Cambridge University Press; 1998.

[8] Engeström Y, Middleton D, editors. Cognition and Communication at Work. Cambridge: Cambridge University Press; 1996.

[9] Wales R, O'Neill J, Mirmalek Z. Ethnography, Customers, and Negotiated Interactions at the Airport. IEEE Distributed Systems Online 2002(September/October):15-23.

[1] [10]Seely Brown J, Duguid P. The Social Life of Information. Boston, Mass.: Harvard Business School Press; 2000.

[2] [11]Bowker GC, Star SL. Sorting Things Out: Classification and Its Consequences. Cambridge, Mass.: MIT Press; 2000.

Address for correspondence

School of Information Systems, GPO Box 252, Hobart, TASMANIA, 7000; Phone +613 6226 6200; Fax: +613 6226 6211; Email: Jo.Kelder@utas.edu.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Health Informatics: An Intercultural Perspective

Quynh Lê

Department of Rural Health, University of Tasmania, Australia

Abstract

Health informatics is a significant contribution to health care. It provides health professionals with powerful technologies to enhance their performance in caring for patients. The introduction of health informatics has added a new dimension in the health discourse. However, there are also issues and problems which are associated with health informatics, particularly in relation to privacy, confidentiality and data security, which are deeply embedded in culture. As privacy and confidentiality are linguistically and culturally constructed, health workers, patients and the public may have different views and exhibit different behaviours towards health informatics. The discussion of these issues is situated in an intercultural discourse.

Keywords:

electronic health record, health care systems, data integrity, electronic medical record, health informatics.

Introduction

Computer has permeated many aspects of society. It is hard to imagine any social activities which are immune directly or indirectly from the influence of digital technology. Computer technology is one of the fastest changing technologies. Technologies which were developed several years ago can easily become out-of-date today. In health science and health care, the impact of computer technology is widespread [1-3]. The development of health informatics is indicative of the impact of computer technology in health science and health care.

Health informatics is a timely contribution to health science and health care in Australia. It is an indication of the growing power of computer technology in health science and health care. While health informatics has made many contributions [4-6], there are also problems which have been identified [7-10]. This paper focuses on the cultural dimension of health informatics.

Health informatics: a new paradigm

The impacts of computer technology in society are marked with the prefix 'e' in many areas of computer development and implementation such as e-learning, e-communication, and e-commerce. In health, the broad term 'e-health' covers a wide domain including electronic health records, health information networks, telemedicine services, health portals, etc. It is an umbrella term covering two areas: health informatics (collection, analysis and movement of health information and data to support health care), and telehealth (videoconferencing and website delivery of health information or health care to a recipient). Health informatics is the appropriate and innovative application of the concepts and technologies of the information age to improve health care and health [11].

Health informatics has much to offer in community health care. Computer networks and telecommunications provide particular support that can enhance the collaboration among clinicians, care providers and patients. Special-purpose computer tools referred to as Consumer Health Informatics (CHI) represent the application of computer and information technologies specifically to support the health information and communication needs of patients and lay persons [12]. Health informatics plays an important role in the management of health information, particularly information of patients. It enables healthcare workers and policy makers at different management levels to plan and manage services. For example, health screening planning does not function well if there are no well-kept records of individuals who have undertaken certain kinds of tests or missed them due to personal or service problems. Health informatics may also record information about patients' health care experience, treatment and financial costs.

The Danish health information network MedCom [13] is a good illustration of health informatics implementation. It handles over 80,000 messages daily. All hospitals, pharmacies and emergency doctors, 90% of general practitioners, 98% of laboratories, 55% of specialists, and 20% of municipalities are connected to it. MedCom enables hospitals to use electronic referrals, and avoid data re-entry. The professional quality of referrals has risen, and discharge letters are stored directly [14].

It is worth pointing out that while e-health is becoming powerful tool and can make a huge contribution to health care, it is still at an early developmental stage in many countries [15].

Cultural factors in health informatics

Culture has been defined as the shared products of the society, including the ideas, norms, and material objects that describe how people handle daily tasks and make

sense of their experiences. Culture is also dynamic and adaptive [16].

> The culture of an individual has a profound effect on the perspective from which they deal with health and illness. Culture has influenced peoples' convictions, attitudes, types of knowledge, and values; modes of behaviour, habits and customs; language and tradition.[17]

Acculturation is a process in which people of a different cultural and social discourse have adapted to accommodate a new discourse. It can be a process filled with confusion, resistance, and reluctance and sometimes sufferings. Health informatics is not just a technology or a simple approach which can be introduced to a human discourse without any problems. There are cultural and social issues associated with health informatics.

> We also learnt that when things go wrong – as they seem to do in more than half the cases – people tend to blame 'the technology' whereas social, behavioural, psychological, and cultural factors are the most likely culprits [18].

The development and implementation of health informatics in health care can go through an acculturation process, which may include negative and positive experiences.

The first problem can be expert-orientated. Health informatics experts and enthusiasts can contribute to the formation of negative attitudes among prospective health informatics users. The worst case is when such experts hold the assumptions that health informatics is the magic solution to heath care and do not take the social and cultural factors seriously. Secondly, the introduction of health informatics can be seen as a paradigm shift in certain discourses. According to Roberts [19] this is the rejection of one set of values and ideas and the adoption of a new set with regards to what constitutes effective implementation. This paradigm shift is occurring worldwide but faster in some parts than the others depending on the availability of resources, existing infrastructure and the stage of development reached. If health informatics is viewed as a new paradigm, strategies have to be planned carefully to facilitate acculturation of current and prospective users to a new health care discourse. Otherwise the acculturation experiences can be painful and sometimes destructive. It is important to involve users (e.g. doctors, nurses, patients, administrative staff, etc) in the decision making process in their acculturation into an unknown or less familiar territory.

Health informatics operates under key principles covering confidentiality, privacy and security. These three concepts are inter-related and are important in evaluating the success or otherwise of health informatics implementation. However, concepts and principles such as privacy, confidentiality and security which govern health informatics have different cultural meanings and values and they are perceived differently by users of different cultural backgrounds. Thus, these three fundamental concepts and principles in health informatics should be examined in terms of cultural discourse.

The cultural discourse of privacy

Humans are social beings. Individuals live together in a community. They belong to a community but this does not mean that their community owns them. They have the right to be left alone. Individuals are entitled to personal privacy which covers three domains:

- *Physical privacy*: such as bag searching, use of our DNA
- *Information privacy*: the way in which governments or organisations handle our personal information such as our age, address, sexual preference and so on.
- *Freedom from excessive surveillance*: our right to go about our daily lives without being monitored or have our actions caught on camera. [14]

Health informatics should adhere to the privacy principle to ensure that individuals' privacy is respected. We tend to take information privacy for granted or do not seriously appreciate it unless it is threatened or lost. Individuals' health information is their personal privacy which should not be 'violated' by government agencies. In special cases when individuals' health condition is a serious threat to the community, their right to privacy may be exercised differently. For example travellers contracted a highly contagious life-threatening disease are expected to reveal fully their conditions to health authorities.

According to Le [20], privacy is something which is personal, belonging to an individual and is not in the public domain. It normally refers to an individual's private life. Thus, according to this definition, an individual's life consists of private and public domains. The private domain includes his/her personal belongings such as home, relationship, thoughts and feelings. The public domain includes social belongings such as professional life, policy, social activities. The following example illustrates what information is private and what is public.

> Mr. Green is working for a company in Tasmania. He joined the Liberal party when he was a student and now he is an independent. His mother is very poor and old but Mr. Green seldom visits his mother even though they are living in the same suburb. They argue a lot when he visits her.

The text given above consists of two kinds of information: private and public. The problem is that the text does not linguistically mark the information in such a dichotomy. To a great extent, privacy is culturally determined. What is private to an Australian may not be so to a Vietnamese.

Not all cultures view privacy in the same way. In Western cultures, individuality is very important. Each person is entitled to their own privacy. Children are introduced to the concept of privacy at an early stage in their childhood. They are taught to respect other people's privacy and they also expect others to respect theirs. In Asian cultures, the division between the public domain and the privacy of individuals is not always clearly prescribed.

In a report about an intercultural experience of a group of Australian students in Australia, Harbon [21] described an

instance in which an Australian student was very upset when she discovered that her host family had searched her suitcase while she was billeted by them. To her it was a serious violation of privacy. Whereas, the host family felt it was interesting to know more about their guest, whom they treasured and cared for tremendously.

Collectivism is very strong in Asian cultures. In an Asian family, privacy is not greatly valued. Parents have 'the right' and 'the duty' to know the private life of their children. It is not a matter of privacy intrusion but a responsibility of the parents to know their children' private domain well so that they can adequately and meaningfully protect their children and ensure their wellbeing. In a Confucian society, interpersonal relationship is the foundation of social coherence. This relationship is characterised by the social roles assigned to each member in a family and in a community. While it is a social violation to ask personal questions in Western cultures, it is a common speech subject in many Asian countries to inquire about someone's age, health conditions, and personal life.

Privacy is an important factor in health informatics. However, users of health informatics may interpret this concept differently due to their social and cultural backgrounds. It is possible that migrants in Australia may violate the principle of privacy in health informatics without being aware of the seriousness.

The cultural discourse of confidentiality

Confidentiality refers to the treatment of information disclosed or provided by individuals on the basis of trust that it will not be made available or disclosed to unauthorised people or services. In health, generally the patient's consent must be sought before his information can be used for a specific purpose.

According to the Australian National Privacy Principles [22, 23], an organisation must take reasonable steps to protect the personal information it holds from misuse and loss and from unauthorised access, modification or disclosure. It must take reasonable steps to destroy or permanently de-identify personal information if it is no longer needed for any purpose for which the information may be used or disclosed.

A study conducted by Lindenthal, Thomas, and Ghali [24] compares the handling of confidentiality among American, Egyptian, and Israeli psychiatrists, and American and Israeli psychologists and internists. The study supports the view that no significant differences exist between practitioners of the same professional groups practicing in different countries while also showing significant and parallel between-group differences. According to Akhter [25], in some cultures, on one hand, sharing personal information among family members indicates a strong bond of coexistence and on the other hand the desire to keep any weaknesses, medical or otherwise, from the extended family is not uncommon. For a societal structure in which the family plays a central role, both allegiance to the family and a desire to keep its reputation strong is an understandable concept. The bond in an extended family provides solace and support in times of need. However, it can also become oppressive and limiting individual freedom.

From the professional duty perspective, confidentiality is based on the trust between patients and health professionals. McClelland and Thomas [26] suggest that confidentiality is grounded in the principle of respect for autonomy – health professionals explicitly or implicitly indicate to their patients that they will keep confidential the information provided to them. Patients are reluctant to share their private and sensitive information if this trust is lost. McClelland and Thomas point out that the duty of confidentiality exists within a wider social context in which other moral obligations may compete. These competing appeals set limits to medical confidentiality and arise from two principal sources: the patient's best interests and public interest. Problems arise when the patients' best interests vary according to their cultural and religious backgrounds, which may not be easily detected or decided by those involved.

Tai and Lin [27] give an interesting example about the cultural concept and practice of confidentiality in a Confucian society. When a patient has been diagnosed with terminal cancer, the first person to be notified is often not the patient himself, but the head of the family, such as the father or the husband. He then will confer with other family members to see what course must be taken. After the decision is made, the patient may be advised in a disguised way, to ease his anxiety. Furthermore, when considering different treatment options, the family members, especially husband or father, are again consulted first rather than the patient himself/herself. When the patient is a father or husband, the family member who becomes the spokesperson for the family, with whom physicians consult, is usually the eldest son.

Gossiping is a good example of cultural variation in dealing with personal privacy and confidentiality. Quite contrary to the principle of confidentiality, gossiping is a sociolinguistic activity which is widespread among cultures. A gossip is a casual conversation between at least two participants about the private life of someone. Morally it is an offence to participate in gossiping. However the seriousness of this moral offence is perceived differently in various cultures. In Western societies, gossiping is condemned and it could be treated as a criminal act if it is proved to cause damage and harm to the victim. In Asian cultures, gossiping is generally discouraged but it is not treated seriously. Gossips are often mentioned in folktales and historical events. The acceptable attitude towards gossiping is a big concern to health informatics as it violates the principle of confidentiality as health workers are expected by the health authorities, patients and the public to strictly adhere to this principle and they should incorporate the spirit of Hippocratic Oath into the social contract.

The cultural discourse of security

To protect individuals' privacy and confidentiality, it is important to ensure that security measures are taken so that health data is kept safely. In health informatics, computer technology provides a range of approaches and strategies

to improve security of health data. Two main approaches include restriction of access and anonymisation of records. Security protection of data requires sound physical as well as logical access controls. Encryption is a method for anonymising electronically held patient information. It is the process by which data are converted into a sequence of alternative characters, by applying a set of rules (or keys) that both generates the encrypted material and is capable of recreating the original information. Another method for anonymising patient information is the use of separate databases in which clinical information is separated from patient-identifier information. The secondary database retains the non-identifiable patient information, which may be used for a range of purposes[28].

Security is an important factor in health informatics. The loss of or unauthorised access to personal and sensitive data can result in financial and legal costs and personal trauma. From an intercultural perspective, there are two issues involved. Firstly, health workers and patients of different cultural backgrounds may treat data security differently. Data security in health informatics needs absolute commitment from those who are privileged to have authorised access. However, such commitment can vary due to different cultural attitudes towards data security and the cultural discourse in which security is reinforced.

One of the most common computer security problems is the management of passwords. A password is a key to access a computer system or a computer file. Though technical security can be very effective, it is the user whose handling of passwords can make computer security vulnerable. It first appears that culture has nothing to do with password security. However, human errors reflect cultural influence on users' attitudes and behaviours in dealing with computer security. In a culture which emphasises collectivism, sharing is a common feature in human interaction, particularly among family members and close friends. Ownership does not belong to individuals but it can be extended to close others. Friendship and kinship are based on mutual trust. In this cultural context, sharing security passwords can occur. Health informatics should take into account this cultural phenomenon.

Implications for policy planning

According to McClelland and Thomas [26] there is a need to establish a new culture for handling health care information – a culture that recognises, understands and responds to the changing structure of health care and health care delivery systems, which depends increasingly on the ready sharing and manipulation of patient information. The digitised health communication and interaction has not only provided an innovative approach to health care but also created a new discourse of health care which requires adjustment and adaptation. Policy makers and health authorities need to introduce programs and strategies for health workers to facilitate their acculturation into the new digitised health discourse.

Australia is a land of cultural diversity. Health workers and patients come from different cultural backgrounds, which may affect their behaviours and attitudes towards health issues and health care, particularly in relation to privacy, confidentiality and data security. As privacy and confidentiality are linguistically and culturally constructed, one would expect different views and behaviours of health workers, patients and the public in response to health policy.

The Linguistic Relativity Hypothesis [29] states that language is so intricately linked to its own culture that it is impossible to fully understand the message through a different language.

> The 'real world' is to a large extent unconsciously built on the language habits of the group. We see and hear and otherwise experience very largely as we do because the language habits of our community predispose certain choices of interpretation (p.177).

The implication for policy planning is that we should not assume that lucid translation of written and spoken health information from English to other languages or vice versa automatically leads to perfect understanding and interpretation.

Cultural diversity should be taken into account when developing and implementing health informatics programs that reflect culturally and linguistically diverse population [30]. Miscommunication or communication failure in the health discourse tends to happen to migrants whose knowledge of English is very limited or whose cultural metaphors and stereotypes influence their health behaviours and attitudes.

Conclusion

In summary, this paper has discussed some cultural issues associated with health informatics. The focus is on issues relating to privacy, confidentiality and security which are fundamental in the implementation of health informatics, particularly from an intercultural perspective.

Health informatics is a significant contribution of computer technology to health care. Metaphorically it is like a superhighway which traverses various roads and alleys of the health discourse, locally, nationally and globally. It has enabled health professionals and health services to improve their effectiveness. However, it is not all smooth. In a culturally diversified discourse, the implementation of technology in dealing with people needs to take into account the social and cultural aspects of human behaviours and attitudes. It is no exception with health informatics, particularly in Australia, which is a land of cultural diversity.

Acknowledgments

I would like to thank Cecilia Chiu for her valuable proofreading and comments.

References

[1] Ammenwerth E, Iller C, Mahler C. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. BMC Med Inform Decis Mak 2006;6:3.

[2] Curry A, Knowles G. Strategic information management in health care - myth or reality? Health Services Management Research [NLM - MEDLINE] 2005;18(1):53.
[3] Dixon DR. The behavioral side of information technology. Int J Med Inform 1999;56(1-3):117-23.
[4] France E. Data privacy in medicine. A perspective offered by the Data Protection Registrar. British Journal of Healthcare, Computing and Information Management 1997;14:20-2.
[5] Lewis A. Health informatics: information and communication. Advances in Psychiatric Treatment 2002;8:165–71.
[6] Linda K. An electronic emergency. Modern Healthcare 2003;33(19):21.
[7] Beauchamp T, Childress J. Principles of Biomedical Ethics. Oxford: Oxford University Press; 1994.
[8] Binetti P. Biotechnology and the birth of a third culture. Journal of Biological Regulators and Homeostatic Agents 2004;18(3-4):255-60.
[9] Brooks F, Scott P. Knowledge work in nursing and midwifery: an evaluation through computer-mediated communication. Int J Nurs Stud 2006;43(1):83-97.
[10] MacPherson S. Health informatics and the cultural divide: a UK perspective. Health Informatics Journal 2001;7(2):62-5.
[11] Health Informatics Society of Australia (HISA). What is health informatics? [cited 2006 18 October]; Available from: http://www.hisa.org.au/100127.php.
[12] Brennan P. Health informatics and community health: support for patients as collaborators in care [cited 2006 8 October]; Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=10805012&dopt=Citation.
[13] MedCom. MedCom projects. MedCom; 1999-2006 [cited 2006 18 october]; Available from: http://www.medcom.dk/wm109991.
[14] Lawlink NSW. What is privacy? The NSW Privacy Commissioner; 2004 [cited 2006 18 October]; Available from: http://europa.eu.int/eur-lex/pri/en/dpi/cnc/doc/2004/com2004_0356en01.doc.
[15] Commission of the Europe Communities. e-Health - making healthcare better for European citizens: An action plan for a European e-Health Area; 2004 [cited 2006 18 October]; Available from: http://europa.eu.int/eur-lex/pri/en/dpi/cnc/doc/2004/com2004_0356en01.doc.
[16] Carolyn P. The impact of culture on rural women's descriptions of health. Journal of Multicultural Nursing & Health 2001;7(1):50.
[17] Janelle T, Celeste Mulry B. Palliative Care and Culture: An Optimistic View. Journal of Multicultural Nursing & Health 2006;12(2):28.
[18] Jadad A, Delamothe T. What next for electronic communication and health care? BMJ 2004;328:1143-4.
[19] Roberts D. The educational paradigm shift: possible implications for higher education and flexible learning. In: Osborne J, Roberts D, Walker J, editors. Open, flexible and Distance Learning: Education and Training in the 21st Century. Launceston: University of Tasmania; 1997.
[20] Le M. Privacy: an intercultural perspective. In: Lakoff R, Ide S, editors. Broadening the Horizon of Linguistics Politeness. Amsterdam: Johnh Benjamins; 2005.
[21] Harbon L. Interpreting culture: preservice teachers taking control and making meaning across cultures. Language, Society and Culture [serial online] 1999 [cited 10 October 2006](4):5-12. Available from: http://www.educ.utas.edu.au/users/tle/JOURNAL/.
[22] The Office of the Privacy Commissioner - Australia. National Privacy Principles (Extracted from the Privacy Amendment (Private Sector) Act 2000). Privacy Commissioner, 2000 [cited 2006 10 October]; Available from: http://www.privacy.gov.au/publications/npps01.html.
[23] Inan H. Australian National Privacy Principles by Hurol Inan 2001-2006 [cited 2006 10 October]; Available from: http://www.hurolinan.com/book/br_detail.asp?LocatorCode=201.
[24] Lindenthal J,Thomas C, Ghali A. A cross-cultural study of confidentiality. Social Psychiatry and Psychiatric Epidemiology 1985;20(3):107-51.
[25] Akhter J. Case Study: Confidentiality in an Extended Family; 2003 [cited 2006 12 October]; Available from: http://www.parkridgecenter.org/Page1885.html.
[26] McClelland R, Thomas V. Confidentiality and security of clinical information in mental health practice. Advances in Psychiatric Treatment 2002;8:291-6.
[27] Tai M, Lin C. Developing a culturally relevant bioethics for Asian people. Journal of Medical Ethics 2001;27:51-4.
[28] NHS Connecting for Health. Better information for health, where and when it's needed; 2005 [cited 2006 18 October]; Available from: http://www.connectingforhealth.nhs.uk/.
[29] Littlejohn S. Theories of Human Communication. New Mexico: Wadsworth; 2002.
[30] Kreuter MW, McClure SM. The role of culture in health communication. Annual Review of Public Health [NLM - MEDLINE] 2004;25:439.

Address for correspondence

Dr Quynh Lê
Department of Rural Health, University of Tasmania
Locked Bag 1372, Launceston, Tasmania Australia
Tel: +61 (0)3 6324 4053; Fax: +61 (0)3 6324 4040;
Email: Quynh.Le@utas.edu.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Successful Systems Sustaining Change

Sheila Bullas[a], John Bryant[b]

[a] *Associate Director, Princess Alexandra Hospital NHS Trust, UK*
[b] *Fellow, Centre for Health Informatics Research and Development (CHIRAD), UK*

Abstract

Much has been published on the success and particularly the failure of IT projects; still failures are commonplace. This prospective study focused from the outset on assessing risk of failure and addressing critical success factors. The aim was to apply existing methods in a challenging acute care hospital where success demanded rapid achievement of sustainable improvements in clinical and administrative processes. The implementations were part of the English National Programme for IT. The desired outcomes required the integration of accepted tools and techniques to provide a pragmatic approach to systems implementation: Lean, Six Sigma, PRINCE2 and Benefits Management. The outcome and further insights into success and failure of IT projects in healthcare are described. In particular lessons are identified related to the business need for the project and the successful achievement of the required benefits and business change.

Keywords:

information systems success and failure; implementation strategies; organizational change

Introduction

The organisation

The Princess Alexandra NHS Hospital Trust (PAH) is a 550 bed general acute hospital. This previously failing hospital had made great improvements over the previous two years; a new senior management team was in place. The organisation required further improvements to reach the standards defined by the Department of Health and other external stakeholders. Information systems were poor, basic information governance was lacking and there were few informatics procedures in place. As in most NHS hospitals there was immense unrealised talent amongst the workforce. The strategy and systems offered by the national programme suited this hospital well. This project started in October 2004.

In common with many NHS organisations, the Trust was under considerable financial pressure, being required to pay back debts and make efficiency improvements.

The English National Programme for IT (NPfIT)

The programme originated from the 1998 Department of Health Strategy entitled 'Information for Health' [1]. A supporting document [2] outlined the information and IT system required for delivery of the NHS Plan [3] and to support patient centered services.

'Securing our Future Health' [4] concluded that to meet people's expectations and deliver high quality the UK needed to devote more resources to healthcare matched with reform. Information and communications technologies (ICT) were recognised as a major driver of this reform. NPfIT was formally established in October 2002 to procure, develop and implement modern, integrated IT infrastructure and systems for all NHS organisations in England by 2010.

NPfIT is a wide ranging programme covering national infrastructure and applications as well as applications to support local organisations. This project focuses on applications to support clinical care and administration of the acute hospital: the incremental implementation of an electronic patient record. The systems and services of a Local Service Provider (LSP) were procured nationally.

NHS Connecting for Health is an agency of the Department of Health whose purpose is to deliver the National Programme for IT.

Local implementation at the Princess Alexandra Hospital

The local implementation comprised the replacement of the existing patient administration system (PAS) with a strategic PAS based on which clinical functionality would be built.

In 2004, the Trust initiated the first project of the programme, namely the replacement patient administration system, the implementation of order communications and the implementation of a data warehouse. In 2005, a second project commenced to replace the radiology information system (RIS) in preparation for the implementation of the picture archiving and communication system (PACS).

Accenture was the Trust's Local Service Provider (LSP). The systems being implemented were iSoft (originally intended to be Lorenzo and later changed to the more established ipm and icm products), HSS (RIS) and AGFA (PACS).

Approaches and methods

Lessons on success and failure

A great deal has been written on the success and failure of information systems. The definition by which success of this project would be measured was adapted from the work of Robert Block in 1983 [5]. Success of the project was defined as: the implementation of the systems on time and within budget, meeting their goals and specified requirements and satisfying the users.

The risks of failure for this project were distilled from sources [6-9] that addressed failure related to information systems implementations. These were used to identify the critical success factors for the project:

CSF 1. Implementation of systems that were tried and tested.

CSF 2. Securing and retaining the commitment of the Trust Board and Executive.

CSF 3. Senior management having limited patience when waiting for tangible results or when delays occur.

CSF 4. Describing, through the project objectives, how the project contributes to the main aims and objectives of the organisation.

CSF 5. Providing sufficient and appropriate resources (people and money)

CSF 6. Recognising that the project is principally a business change project with IT as a major enabler of that change. Linked to this the acknowledgement that change must come from within the organisation and may be facilitated from outside.

CSF 7. Ensuring robust management of the project, including project governance.

CSF 8. Understanding, influencing and managing the expectations of users.

CSF 9. Communication. Linked to this, recognising the fear of retribution when conveying bad news to the Executive and Board.

In addition, the design took into account the view that most failures are the victims of organisational and people related issues [10], even those which initially appear as technical failures. The point at which the failure becomes inevitable may be many months before failure becomes apparent. This suggests that avoiding failure should be part of project design.

There were a number of constraints outside the control of the Trust within which it had to operate:

- Governance arrangements from the Trust, through the local health community, the Strategic Health Authority, to Connecting for Health. This defined the structure of Project and Programme Boards, structure and content of key documentation and project dates.
- The systems and the suppliers were determined through the national procurement. This gave a good fit with the local strategy for improvement.

Approaches, tools and techniques

A range of tried and tested approaches, tools and techniques were adopted for the project: some at the outset and others emerged and were integrated during the project as a result of changes in the internal and external environment. The choice was not an entirely free one: it included best practice standards, approaches required by NPfIT, and those adopted for the Clinical Service Improvement programme [11].

Project management

PRINCE 2 is the UK Office of Government Commerce (OGC) recommended approach to project management [12]. The approach was tailored to suit the nature and size of the project and to fit with requirements of the LSP and reporting bodies. Once the approach was agreed, this was strictly adhered to. Within the Trust, this standard was poorly implemented before starting this project, even for IT projects where it is a requirement. PRINCE2 was used for both technical implementation and business change. A qualified and experienced PRINCE 2 practitioner was an essential requirement of the project manager.

Change management

Change management was entirely the responsibility of the Trust. The LSP introduced a Business Change Workstream and supported the Trust in undertaking Business Change Workshops to a specific format. However LSP business change covered only the simplest level of change where the new system necessitated the redefinition of processes at the systems use level.

It was necessary to adopt additional methods if business change was to be created and the objectives of the project achieved. The following model shown in Figure 1 was designed, describing the different levels of change that were required if the objectives of the system were to be met. This was developed from various concepts of change management [13,14].

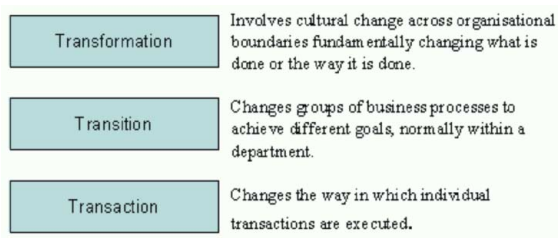


Figure 1 - Levels of change

These different levels of change required different approaches.

At the **transaction level**, workshops introduced by the LSP were used. In discussion with the users, the current 'as is' processes for all transactions involving the old and/or new information systems were documented. Users worked with the new systems in order that they could envisage how transactions would change ('to be' processes). These workshops also stimulated ideas for transitional change by encouraging discussion on the effi-

ciency and effectiveness of current processes. At the transaction level existing transactions or processes (series of transactions) are replicated using the new system. These new processes were documented and used to tailor the training which focused on business processes and not just what buttons to press. While it is essential that all current processes are comprehensively translated to the new system, to only do this gives rise to limited benefits and would not have met the objectives of the project.

During the transaction level workshops, areas were explored where transition level change might be identified. This included identifying bottlenecks and constraints in the processes, elements of waste and inefficiency, and duplication. These ideas input to the **transition level.** Further changes were identified at this level by conducting benefits workshops using the Cranfield Benefits Management method (see below). Processes were re-designed in workshops involving all stakeholders in those processes. Re-designed processes, including changes to who did what, when and where, were documented and tested before individual transactions were defined. The aim was to optimise whole processes including both computer based and manual steps.

The Trust was subject to more fundamental change as a result of external requirements, for example the introduction of the 18 week referral to treatment target [15], and internal pressures to improve performance, eliminate waste and achieve financial balance. In line with healthcare organisations around the world, NHS organisations were beginning to adopt industrial processes to improve healthcare [16,17]. The Trust adopted a tailored Kaizen approach as the means by which it would achieve **transformational change.** This approach is based on Lean Thinking, started by Toyota in the 1950s and developed by Womack and Jones in the 1990s [18]. The approach focused on the patient journey. It aimed to ensure that processes flow efficiently, value to the patient is increased, and steps that fail to add value are minimised. This approach provided a single transformational change programme within the hospital. The information systems implementations and improvements needed to integrate with this programme if maximum success were to be derived from both the transformation and the systems implementation programmes.

Benefits management

The Cranfield Benefits Management method [19] was adopted by the Trust as its standard to identify the benefits, particularly transitional benefits that would arise from the implementation of the information systems. This approach was adopted by much of the NHS for use as part of the NPfIT. The approach provides a structure to identify benefits that is linked firmly to the business drivers for the project. These drivers are derived from the business strategy of the Trust. In summary, workshops involving all key stakeholder groups systematically identify and measure benefits plus business and IT enablers to achieve the benefits. Benefits are measured and ownership for realisation established at this early stage.

This method was adapted locally to ensure that benefits were linked into the main strategic objectives of the organisation at the outset and further refined at stages throughout the project. Early workshops identified the processes to be re-designed as part of the Transitional Change Plan and provided clear guidance for configuration and training work packages. Output from this stage provided the means of benefits monitoring and the basis for benefits realisation.

Stakeholder analysis, as included in the Cranfield Benefits Management approach was also the basis for the development of the communications plan. This identified the stakeholders, their current commitment, and that required for success.

Organisation-wide Lean events identified transformational change that existing and new information systems were required to support. The proposed changes were further subjected to benefits workshops to ensure that information flows supported both new and continuing processes.

Resources: people and money

When an organisation is under financial pressure, as this one was, it is tempting to under-estimate resource requirement in order to get the project agreed. This limits the success that is achievable if the project is attempted with inadequate resources. It also undermines the credibility (and hence challenges commitment) if additional resources, that might have been foreseen at the outset, are requested at a later stage. For these reasons, great attention was given to estimating the required resources as soon as possible, ensuring (through project governance structures and processes) that these were clearly understood and then robustly defended. This particularly applied to the people and their skills.

Resource requirements, identified during plan development, led to the development of roles, job descriptions, person specifications and recruitment of the project team.

Results

The first system to go-live (RIS) met the definition for success: on time, within budget and meeting the user's expectations. PACS goes live in May 2007 and has met all early milestones successfully on or before planned dates.

The PAS/Order Communications package has been deferred for two to three years in a complex scenario including supplier delay and local financial pressures. While an option was presented to the Executive that was highly likely to have brought the project in on time and within budget, the delays had compromised confidence and commitment with not entirely unexpected consequences.

Discussion

This section discusses the progress in the light of each of the critical success factors.

The programme being implemented ranged from the tried and tested to the new and not fully developed (CSF1). This

was inextricably linked to CSF3: Executive patience is limited when results are not quickly coming or delays are encountered. Furthermore, different systems were subject to different contractual and financial conditions. This proved important in the progress and outcome of the various elements in the project. Commitment for the PAS/ Order Communications package was lost when delays resulted in the supplier being unable to deliver on time. This was due to delays in the development of the Lorenzo application. Although substituted with tried and tested applications, confidence was already compromised. Commitment might have been sustained had the delay not coincided with financial pressures which led the Trust to withdraw support. On the other hand, the tried and tested RIS and PACS systems, which came with a far higher price tag but with financial penalties on the Trust if milestones were missed, sustained commitment (CSF2). Although it was made clear that support for this could also have been withdrawn had financial arrangements been different. Financial balance in the current year took precedence over all other objectives.

The formal governance arrangements were one essential element in gaining and sustaining the commitment of the Executive and Trust Board: providing formal communications on progress against plans and budgets (CSF9). At least as important were the close working relationships without which continuing support cannot be expected at times of pressure. This was a particularly important aspect of this project: severe financial pressure combined with delay in delivering the technology put strains on the commitment.

How the project objectives contributed to the aims and objectives of the organisation was clearly identified through the business change workstream and communicated to the Executive and Board through project governance arrangements established through the formal project management arrangements (CSF4). Board and Executive understanding of this was maintained throughout.

Provision of sufficient and appropriate resources (people and money) was supported by the Executive (CSF5). In the financial climate, there were, unsurprisingly, attempts to compromise on project resources in order to reduce costs. However the risk that this posed was understood and appropriate resource levels maintained. The relationships developed with executives and the open and robust approach to project governance proved pivotal in maintaining resources.

The programme was managed very much as a business change project but not to the exclusion of excellent technical implementation of the systems: both are essential requirements for success (CSF6). The adoption of the 3-tier model (Figure 1) and associated methods, the inclusion of a business change workstream within the project structures and processes, and the full integration of information change into the transformation programme for the Trust as a whole were crucial to success. The latter proved to be the most difficult to achieve. The Trust-wide transformation programme was defined as separate from informatics and promoted as such by people both within and outside the organisation. As a result, project processes were needed to make up for shortcomings in the level of integration. There can only be one major change programme in an organisation, to which IT projects contribute. In the NHS NPfIT, conflicts could arise as a result of different external organisations imposing different approaches.

Robust management of the project, including project governance (CSF7) was ensured through the adoption and implementation of PRINCE 2 appropriately tailored to nature of the project. In practice this included defining governance structures, regular meetings, appropriate agenda, maintained risks and issues logs, comprehensive status reports, and response to exceptions agreed at the right level. This gave confidence that the project was progressing to plan and any variations were dealt with before critical milestones were in danger of being missed.

Understanding, influencing and managing the expectations of users is at the heart of the very definition of success (CSF8). The business change programme was designed to incorporate this requirement. It is inevitable at the start of a project that the expectations of the users are diverse and contain some elements that the project cannot meet. The workshops brought as many of the users into the process as was practically possible. Developing user understanding of the project and the products brought expectations into line with what could reasonably be delivered and the project was adapted to more closely meet expectations where possible. Meeting those expectations was then dependent on the successful technical and organisational implementation of the project as agreed with the users. One aspect of user expectation management that had not been adequately addressed was at the immediate go-live period. The implementation of the RIS system went smoothly with few teething problems that were promptly and efficiently addressed. Users had not been prepared to expect these and, since all other aspects of the project (including time and budget) were as expected, this caused some temporary dissatisfaction.

The final critical success factor (CSF9) addressed in this study was the need to communicate. The communication plan was an important part of the project but just what constituted appropriate and adequate communications? Stakeholder analysis, as included in the Cranfield Benefits Management approach, was the basis for the development of the communications plan. This determined the minimum requirements for communications. In practice, communications played a greater role than such analysis suggests. The balance between listening and telling, the credibility of the person or medium providing the communication, and the need to continually assess and repeat or reinforce should be emphasized.

Fear of retribution when conveying bad news to the Board, or in this case Executive, was part of the culture and conveying news contrary to the pertaining view of the Executive was not welcomed however diplomatically expressed.

Conclusions

This study set out to determine whether one could take lessons about success and failure from the past, systematically design a project to take account of these lessons, and as a result improve the likelihood of success. From such a case study approach, it is difficult to know conclusively. It is uncertain whether progress or outcomes would have been worse had this not been the case.

This case study has shown that attention to all the critical success factors chosen proved important to the success or failure of the projects. One factor that became apparent during the study was the subjective nature of the critical success factors and how they were interpreted in the design of the project. Just how did you judge whether the commitment of the Executive was sufficient to ensure that it would not disappear under difficult circumstances? How could you define the required relationships required to ensure the engagement of clinicians and other stakeholders?

The critical success factors and how they should be addressed should be further defined for the benefit of those implementing projects. Small organisations cannot expect to develop and retain the expertise necessary to direct the complex scenarios that lead to failure of large IT projects. Expert external resources may provide support but cannot replace a strong internal lead. Given the wide ranging nature of the factors and cultural impact, the informatics professionalism of the local programme director and their direct relationship with the Executive is proposed as an additional critical success factor.

Hospital IT projects aim to create change and achieve benefit within a complex environment. It is necessary to work within the broader constraints and not against them, and to turn these constraints to advantage. This requires adopting the change approaches that the organisation is comfortable with rather than those the IT profession might wish to adopt.

Acknowledgments

Thanks to all the staff at the Princess Alexandra Hospitals NHS Trust and all others involved in the projects including local health community, Strategic Health Authority, Connecting for Health and the suppliers.

References

[1] Department of Health, NHS Executive. Information for Health an Information Strategy for the Modern NHS. Dept of Health, 1998

[2] Department of Health. Building the Information Core: Implementing the NHS Plan. Dept of Health, 2001.

[3] Department of Health. The NHS Plan A plan for investment a plan for reform. Cm 4818-1. Stationery Office, 2000.

[4] Wanless D. Securing our Future Health: Taking a Long Term View. HM Treasury, 2002.

[5] Block R. The Politics of Projects, Yourdon Press, New York, 1983.

[6] Sauer C. Why Information Systems Fail: A Case Study Approach. Alfred Waller, 1993.

[7] Collins T. Crash: ten easy ways to avoid a computer disaster. Simon & Schuster, 1997.

[8] Fortune J, Peters G. Information Systems: Achieving Success by Avoiding Failure. Wiley, 2005.

[9] Lorenzi NNM, Riley RT. Organisational Aspects of Health Informatics. Springer-Verlag, 2004.

[10] Lorenzi NM, Smith JB, Conner, SR, Campion TR. The Success Factor Profile for Clinical Computer Innovation. In: Fieschi M et al (Eds), Medinfo 2004. Amsterdam. IOS Press, 2004.

[11] Rogers H, Silvester K, Copeland J. NHS Modernisation Agency's way to improve health care. BMJ 2004 328: 463 (letter).

[12] Office of Government Commerce. Managing Successful Projects with PRINCE2 Manual. Stationery Office, 2005.

[13] Noehr C and Bernstein K (1989). Can HIS be developed without Organisational Change? In Barber B et al (Eds), Medinfo 89. Amsterdam. North Holland.

[14] ICL. A window on the Future. An ICL Briefing for Management on the Findings of the Management in the 1990s Research Programme, conducted by the MIT Sloane School of Management. ICL, 1990.

[15] Department of Health. Tackling hospital waiting: the 18 week patient pathway. Dept of Health, 2006.

[16] Young T, Brailsford S, Connell C, Davies R, Harper P, Klein JH. Using industrial processes to improve patient care. BMJ 2004; 328: 162-164

[17] Babensky JA, Roe J, Bolton R. Lean Sigma: will it work in healthcare? Journal of Health Information Management, 2005, vol 19, no. 1.

[18] Womack JP, Jones DT. Lean Thinking. London: Simon and Schuster, 1996.

[19] Ward J, Daniel E. Benefits Management: Delivering value from IS and IT investments. John Wiley & Sons Ltd, 2005.

Address for correspondence

Sheila Bullas, BSc MBA, MBCS, CITP,
iBECK, Willsmead, Abington Pigotts, Royston, Hertfordshire, UK, SG8 0SD.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

MUST - A Participatory Method for Designing Sustainable Health IT

Finn Kensing, Hrönn Sigurdardottir, and Arjen Stoop

IT University of Copenhagen, Denmark

Abstract

Several important issues in designing sustainable health IT, such as coherent visions for change and genuine user involvement, are too often neglected or not paid enough attention to in practice. The MUST method addresses the early stages in the design of sustainable IT applications. The method highlights how those issues can be dealt with as it provides practical recommendations in terms of principles, tools and techniques. The method has proven helpful in assisting project participants to focus on and combine issues that are also crucial when designing health IT. MUST has been developed and tested in commercial settings. Here however, we illustrate the method's potential for health IT as it was recently used for the evaluation of a faulty health IT project intended to support shared care in relation to pregnancy.

Keywords:

health informatics, action research, methodology, systems design, shared care, sustainability

Introduction

Several important issues in designing health IT, such as a coherent visions for change, real user involvement, understanding the work that is supposed to be supported by IT, and the way proposed changes have to be incorporated in daily work, are in practice too often neglected or not dealt with properly [1-4]. As a consequence, sustainability of the IT system lacks. By sustainability we refer to the need for a balance between development, use and protection of an organizations' resources (financial foundation, personnel and IT systems). The MUST method we present here has proven to be able to deal with these issues. The method was developed and tested through a longitudinal action research program involving 14 IT projects in public and private companies [5]. Here the method is used for evaluation rather than for design, therefore the paper is of an exploratory character.

The MUST method

MUST [6] consists of four resources to assist project participants in planning, conducting, and evaluating their projects. By resources we stress that we are not dealing with a cook book type of method, instead project participants each time have to decide which method elements to include and how. The resources are: 1) A range of concepts and conceptual frameworks, 2) Four guiding principle, 3) Four phases and 4) Sixteen tools and techniques. We restrict ourselves to presenting the principles in some detail and to simply list relevant tools and techniques for each of them. Further we give the rationale behind the four phases.

The four phases

The method suggests an IT project to be divided into two separate projects [6], a design project and a realization project. In between the two we find the call for tender and contractual negotiations. MUST deals only with the design project. A design project results in one or more coherent visions for change on the basis of which the customer decides which vision, if any, should be pursued through a realization project. This division is especially needed when IT is designed to support communication and cooperation between actors from different professions and organizations. The MUST method suggests an IT design project to be divided into four phases as illustrated in Table 1. The table shows the focus and the result of each phase as well as the decisions supported by each phase.

Table 1 - The four phases of MUST, their focus, the subsequent results, and the decisions they support

Phases	Focus	Results - decisions
The initiation phase; establishing the project.	The scope of the IT design project: time, costs, contents, participants.	Project charter & a plan - deciding the scope and basis of the IT design project.
The alignment phase; understanding management strategies.	Relations between the goals of the IT design project and the organization's business and IT strategy.	Strategic analysis report - selecting work domains for further analysis.
The in-depth analysis phase; understanding current work practices.	The work practices of the selected work domains.	Analysis report & descriptions of work practices - prioritizing goals, problems and needs, and ideas for IT.
The innovation phase; developing visions.	Visions of IT systems and their relations to the organization of work and the users' qualifications. The realization project.	IT design report & mock-ups and prototypes - deciding which visions should be realized and how should this take place.

The four principles

The principles express an overall perspective built into the method while the tools and techniques provide more concrete recommendations on how to conduct various activities. These may be substituted by other tools and techniques, which serve the same needs. Instead, structuring the design process according to the four phases, and striving to live up to the four principles is considered important – as we shall see in the evaluation below.

Principle 1. Coherent visions for change

The result of an IT design project is one or more coherent visions for change in the organization in question and in relations to its environment. The proposed change should meet the organization's revealed goals, needs, and opportunities within its business and IT strategy - which may itself need revisions as part of the project. By a coherent vision we mean that the following three elements of an IT application are designed to support each other: IT systems, work organization, and the qualifications users need to perform their job with the help of the proposed IT systems in the proposed work organization. If management and clinicians don't share the same vision of how to support the strive for better economy, quality, and service for the patients, the intended changes will not occur [7]. Relevant tools and techniques that support this principle include workshops, prototyping, and scenarios.

Principle 2. Genuine user participation

If users contribute solely as informants, it is not considered genuine user participation. Instead this principle prescribes the active participation of end-user representatives influencing the process of design as well as the visions it results in. There are two rationales for this, a pragmatic and a political. The pragmatic argument is that mutual learning between users and IT designers is needed, while the political argument shows a concern for the employees' rights to influence their own working conditions. The political argument may also be justified by the simple fact that e.g. health care professionals have demonstrated that they have the power to block an IT application that they find irrelevant or too cumbersome to use [8, 9]. Relevant tools and techniques that support this principle include user representatives in project groups and steering committees, prototyping, workshops, and hearings.

Principle 3. IT designers' need to experience the users' work practices

Basically, there are three different ways of obtaining new knowledge relevant for an IT design project: Reading about the subject matter, asking knowledgeable people to tell you about it, or allowing yourself to experience the subject matter first hand. This principle argues for the need to include also the latter to really understand the work practices for which you are developing IT support. In recent years a substantial critique has been put forward to the development of health care IT not taking account of the social aspects of health care work [10]. By observing the users while performing their regular business or while trying out prototypes in situations as realistic as possible, IT designers acquire knowledge and understanding of the work and communication processes that need to be supported by the IT system. Relevant tools and techniques that support this principle include observations, in-situ interviews, and thinking aloud experiments.

Principle 4. Anchoring visions for change

This principle focuses on three groups, whose members usually cannot all participate directly in an IT design project: 1) Management, who have the power to decide whether or not the proposed visions will be implemented, 2) Employees and other interested parties, who will either use the IT systems in question, or who will be affected by them, 3) Internal and external people taking care of the technical and organizational realization activities that are required to implement the proposed visions. The principle prescribes that these groups must be informed and involved in various ways to be able to evaluate the consequences of the proposed changes, as seen from each of their perspectives. This needs to occur in time for the project group to incorporate their reactions into the final design proposals. If visions are not anchored with these groups the expected changes are unlikely to take place. Relevant tools and techniques that support this principle include prototyping, scenarios, reviews, and hearings.

The case: Sundhed.dk

To illustrate the MUST method's relevance within Health Informatics we have used it for the evaluation of a web based pregnancy application. The intent of the project was to support communication and collaboration among midwifes, obstetricians, general practitioners, and the pregnant women. Especially this type of health IT has been proven difficult to manage [11].

The pregnancy project was launched as a part of a large Danish public e-health portal initiative in the year 2004. The portal (in Danish Sundhed.dk) provides a framework for electronic communication between the parties involved in the Danish National Health Service and communications with the patients. Furthermore the portal provides information and services for the citizens and for health care professionals. According to the contract, the vendor who won the tender for the portal should also design web applications to support shared care for two widespread diseases. These web applications should be designed in a way so that they would be able to function as a generic model for future web based services. Even though pregnancy is not a disease it was chosen since pregnant women are known to be competent users of the Internet and web based services. In addition at the outset, pregnancy was considered relatively simple to support. The portal has received several national and international rewards for providing excellent service to citizens and health care professionals and outstanding user interface.

Research methods

For a detailed description and evaluation of the research method used to design the MUST method readers are referred to [5]. Instead here we concentrate on the method-

ological considerations as to how the evaluation of the pregnancy project was carried out. The evaluation is here used as the basis for proposing the MUST method as a resource for designing health IT. The main data collection tools were 1) participant observation at project group meeting over a period of eighteen months (documented in notes); 2) eight qualitative interviews (documented in notes and on tape) with project participants (IT specialists, user representatives, and management form both the customer and the vendor side); 3) a full day workshop (documented in a report approved by the participants), and document analysis.

Results

This section draws on a recent evaluation of the pregnancy project, focusing on the process rather than on the resulting product. The evaluation report was approved with minor corrections by all involved in the evaluation and later used as the basis for a reformulation of sundhed.dk's business strategy and for developing future services. In the evaluation below we end each section by listing proposals for improvements. We use the four principles of the MUST method as a conceptual framework for the evaluation of the pregnancy project, thereby also demonstrating the design potential of the method. The argument is that if the project had lived up to these principles it would have had a greater chance to succeed. Unfortunately, in contrast to clinical trials, we are not in a position to isolate a set of factors and conduct an experiment based on which we may claim evidence as to a design method. This is due to the complexity of IT projects and their highly contextual dependencies.

Coherent visions for change

For the main course of the project the participants did not share a common understanding of what they were designing. This applies to the functionality of the pregnancy application - the first element in MUST's understanding of a coherent vision for change. It also applies to the second element – work organization – since the web application's integration to other IT systems remained a battleground for most of the project's lifetime. Therefore it remained unsettled for a very long time how doctors should organize their record taking in relation to their current EPR's and the new web application. Instead the third element – users' need for new qualification – was by all parties considered not to be a problem. Despite the above problems the web application in itself was considered simple. The vendor produced a preliminary draft for a general description of the expected course of events and of the associated systems integration. However, this should be perceived only as a first general project draft to be developed further in close cooperation with the partners involved in the pilot project. Therefore the participants from the two counties that signed up for a pilot project expected to be able to obtain great influence on the design of the web application. This turned out to be a problem since few resources were set aside for experiments. Thus constant discussions took place on whether ideas for improvements were inside or outside the scope of the contract. Like for the web portal, the idea of the pregnancy application is "simply" to provide access to data already available. The participating counties differed on this issue, as they had different prerequisites for participating in the pilot project. The consequence of this was never analyzed, or at the least no consequences were drawn on the basis of this. The one county did not have an IT system that could be used for generating data for the web application, while the other had such systems. Therefore the participants from the former kept arguing that the web application should support doctors in entering and retrieving health data about their patients, while participants from the latter preferred to use their existing IT systems. The solution became to support both. This landslide in the conceptualization of the product, developed slowly. Some key actors let it happen – they wanted to keep decisions open as long as possible, thus maintaining the reasons for the counties to be part of the project. Others pursued openings to have the application go in the directions they wanted.

The character of the contract between the customer and the supplier was the main reason for the participants not being able to agree on the product's functionality since it consisted of an incomplete requirement specification. Both the customer and the supplier used the contract against the counties in the latter's attempt to get what they wanted. There are both negative and positive consequences of this course of events. The pilot project did clarify to some degree the concept of 'shared care' and how IT may support this. This in turn helped clarifying what 'a generic model for future web based services' might mean. This is crucial for future sundhed.dk projects. On the other hand 'generic' also turned out to be a fuzzy concept. Concerns were raised about the possibility to rely on a generic model in future shared care projects. Finally, the constant questioning of the scope of the project resulted in the inclusion and exclusion of processes and work practices in ways that did not contribute to the progression of the project. The following ideas were suggested based on the evaluation:

- Sundhed.dk needs to take on the job of explaining to its funding partners that its different projects move in unreclaimed territory, which means that room for experiments and systematic evaluations are needed.
- Before commitment to a fixed time, fixed functionality, fixed prize type of contract, the project should have developed a coherent vision for change through an experimental approach using MUST.
- Sundhed.dk has to take the role of coordinating the counties needs for IT support for shared care, or it must see to that another agent plays that role.
- The funding partners of sundhed.dk need to make an informed choice - as to various types of projects - about the role(s) of sundhed.dk at a national level.
- The partners need to understand that different roles come with different prerequisites in terms of project funding capabilities and different qualifications and authority to run a project.

Genuine user participation

Three central groups of users were involved too late and not in adequate ways for them to influence the process of design as well as the product it resulted in. These groups

are: Pregnant women, visiting nurses, and general practitioners and their IT vendors. They are a mix of individual and organizations that are not part of the same line of command. Here there is only room for including an evaluation in relation to the pregnant women.

Pregnant women were involved too late and mainly for testing purposes. The main reason for this was that the project group decided that user needs had been collected once and for all three years before the project really started. However, not even these needs were taken care of. The consequence was that the wishes, needs, and requirements of perhaps the primary user group were never really dealt with. The following ideas were suggested based on the evaluation:

- A proper stakeholder analysis should be conducted before requirements are developed and decided upon.
- Users should be involved in order to develop their wishes, needs, and requirements as well as in an evaluation of the degree to which their interests are met by an application under development.
- An iterative process is recommended for projects where wishes, needs, and requirements are not clearly stated at the outset or where conflicts of interests are to be expected.

Experience the users' work practice

The project group totally ignored this principle. The reason was that they were either afraid to - or not qualified to - expose themselves to the situations for which their IT application was intended. Thus, they managed to keep a distance to the real needs of the various types of intended users. The consequence was devastating for the quality of the pregnancy record, as its functionality and modes of interactions did not meet the needs of its intended users.

The supplier's suggestions for how to meet the requests of the customer, was based on a too narrow understanding of the work processes that the project group finally tried to support. This flaw was never faced. Sundhed.dk took on the task of a usability study, but pregnant women were not included even though an important part of the rhetoric around sundhed.dk was that the portal was about supporting patients/citizens. Further, the clinical work practices were not analyzed.

The following ideas were suggested based on the evaluation:

- Members of the project group should experience at the least one visit of a pregnant woman to her general practitioner's consultation.
- Members of the project group should observe a pregnant woman in her home for one day in the beginning, half way trough, at the end of her pregnancy, and during the first weeks after the delivery.

Anchoring

The project was organized so it should have been possible to live up to this principle. At the national level both management and the vendor, who took care of the technical implementation, were fully able to evaluate the consequences of the proposed changes, as seen from each of their perspectives. Further, both counties had representatives in the national project group and they each sat up county-based project groups. They arranged meetings to inform the clinicians, who were also involved in pilot testing and teaching workshops for them to familiarize themselves with the pregnancy application and provide their feedback. However, one county withdraw from the project after waiting many months for its vendor to estimate the costs of integrating the web application with its local system – the costs were considered too high. The other county also withdrew from the project since the application disappointed the clinicians.

The anchoring problem was not so much due to bad project organization and management at the national or at the county level. Rather, the problems were caused by the lack of a coherent vision for change as already accounted for and by the fixed time, fixed functionality, fixed prize type of contract that did not allow for experimentation and learning.

The following ideas were suggested based on the evaluation:

- Distinguish between at least two types of projects that correspond to different types of contracts. For the one type it is possible to freeze a requirement specification before the start of the technical and organizational realization project. For the other type, user representatives need to take part in iterative explorations of problems and possibilities along side the development of solutions. For the first type of project, a fixed time, fixed functionality, fixed prize type of contract may work. Instead for the other type of project a development contract is needed, that allows for experimentation and learning.
- Establish mechanisms to detect if the wrong project model has been selected, and adjust accordingly.

Discussion

Many participatory design *tools* and *techniques* have been developed [12]. Developing a coherent participatory design *method* has not been the aim of the PD community. However, a few groups have systematically organized their design practices into a coherent method. Most of these are only documented in one ore more papers see e.g. Grønbæk et al [13] and Blomberg et al. [14]. In addition to the MUST method, the only other coherent method published as a usable guideline is Contextual Design (CD). Readers are referred to [15] for a proper comparison of MUST and CD, here we just list that 1) CD and MUST have the same *application area*: a participatory design approach that focuses on early design activities. 2) CD blurs the distinction between managers and users referring simply to 'customer' where as to MUST 'users' and 'managers' are different categories. 3) MUST acknowledges that there may be different interests at stake and it provides ways to deal with them without assuming that conflicts may be resolve in a harmonious way, instead CD is not explicit about conflicts and thus provides no means for dealing with them.

4) MUST provides tools and techniques for project management, CD does not. Other traditional software development methods also have addressed the issues and concerns that are dealt with by the MUST method. There are thought significant differences between the MUST and the traditional methods that we in the following will address shortly. Compared to traditional waterfall models and evolutionary models, MUST is designed to incorporate the best – and avoid the worst – aspects of these approaches. We recommend iteration between analysis and design activities within and among the project phases. Other contemporary methods, like Rapid Application Development (RAD) and Extreme Programming (EP) also have abandoned the traditional phase model, but differ from the MUST method particularly in relation to how and when the decision for what kind of a system is to be built, is acquired. Both RAD and EP have incorporated the assumption that a decision to build a system of a particular kind already has been made. RAD, EP and and Rational Unified Process (RUP) focus on building systmes from scrach. In turn, RUP does incorporate early design activities but has strong focus on modeling, specifications and implementation. Our last comparison is to the Business Process Reengineering method. Beeing simliar in scope, MUST and BPR both aim at formulating visions for the future use of IT and leave out the organizational implementation but consider the relationship between a design project and an organizations business and IT strategies. BPR does not deal with ethical or practical isssues relating to the users and therefore does not provide any help in understanding, developing or presenting relations between IT and user's work practices.

Here we have used elements of the MUST method for the evaluation of a recent failing health IT project. We demonstrated that if two of the method's four resources had been applied, the project would have had a better chance of success. Due to space limitations we did not find room for neither the introduction nor the application of MUST's other two resources. Anyhow, we have demonstrated that methods like MUST are relevant for health informatics especially when it comes to the type of IT applications that have the ambition to support shared care. This type of application is characterized by the fact that neither health professionals nor IT specialist have good models for what the application should be like. Further, such projects need to be conducted in a context in which there is no unity of command. This implies that experiments are called for, and that all parties are allowed to systematically evaluate the degree upon which their interest are taken care of. MUST is such a method.

Acknowledgements

We want to thank Keld Bødker and Jesper Simonsen, who together with the first author of this paper developed the MUST method. Further, the project group of the pregnancy project and sundhed.dk as an organization participated openly and honestly in the evaluation of a rather troubled project for the purpose of improving their performance in future projects.

References

[1] Atkinson CJ, Peel VJ. Transforming a hospital through growing, not building, an electronic patient record system. Methods Inf Med 1998; 37(3): 285-93.

[2] Ash JS, Berg M, Coiera E. Some Unintended Consequences of Information Technology in Health Care: The Nature of Patient Care Information System-related Errors. J Am Med Inform Assoc 2004; 11(2): 104-12.

[3] Berg M. The search for synergy: interrelating medical work and patient care information systems. Methods Inf Med 2003; 42(4): 337-44.

[4] Forsythe DE, Buchanan BG. Broadening our approach to evaluating medical information systems. Proc Annu Symp Comput Appl Med Care 1991: 8-12.

[5] Kensing F. Methods and Practices in Participatory Design [Dr. Scient dissertation]. Copenhagen: ITU Press; 2003.

[6] Bødker K, Kensing, Finn., Simonsen, J. Participatory IT Design. Designing for Business and Workplace Realities: MIT Press; 2004.

[7] Nikula R. Why Implementing EPR's Does Not Bring about Organizational Changes - A Qualitative Approach. Studies in Health Technology and Informatics 2001; 84:666-9.

[8] Levitt JI. Why physicians continue to reject the computerized medical record. Minn Med 1994; 77(8): 17-21.

[9] Scott JT, Rundall TG, Vogt TM, Hsu J. Kaiser Permanente's experience of implementing an electronic medical record: a qualitative study. BMJ 2005; 331(7528): 1313-6.

[10] Coiera EW, Jayasuriya RA, Hardy J, Bannan A, Thorpe ME. Communication loads on clinical staff in the emergency department. Med J Aust 2002; 176(9): 415-8.

[11] Berg M. Health Information Management. Integrating Information and Communication Technology in Health Care Work: Routledge; 2004.

[12] Kensing F, Blomberg J. Participatory Design: Issues and Concerns. Computer Supported Cooperative Work (CSCW) 1998; 7(3): 167-85.

[13] Grønbæk K, Kyng M, Mogensen P. Toward a cooperative experimental system development approach. In: Computers and design in context: MIT Press, 1997; pp. 201-38.

[14] Blomberg J, Suchman L, Trigg RH. Reflections on a Work-Oriented Design Project. Human-Computer Interaction 1996; 11(3): 237-65.

[15] Beyer H, Holtzblatt K. Contextual Design: Defining Customer-centered Systems: Kaufmann; 1998.

Address for correspondences

Finn Kensing,
IT University of Copenhagen, Rued Langgaards Vej 7
DK-2300 Copenhagen S, Denmark, e-mail: kensing@itu.dk

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Change Management and the Sustainability of Health ICT Projects

Karen Day[a] and Tony Norris[b]

[a] *School of Population Health, University of Auckland, New Zealand*
[b] *Institute of Information and Mathematics Sciences, Massey University, Auckland, New Zealand*

Abstract

The development of the electronic health record (EHR) is a strategic and important enabler of the delivery of integrated healthcare. As each innovative aspect of the EHR is implemented in New Zealand, long-term success is essential for its overall sustainability on the national scale. How we achieve this success is dependent upon how people adapt to the changes brought about by the implementation of these innovations. The transition period of the change process we follow during this adaptation is characterized by a capability crisis, in which we tend to predict failure in our attempts to make the changes to which we are committed. This could be a signal of the first step toward sustainable change as people adapt to changed processes, technology and relationships. Once we have mastered the incremental changes brought about by health ICT projects for the implementation of the EHR, we are able to connect health services by means of the same EHR and provide enabled, sustainable integrated healthcare.

Keywords:

change, sustainability, failure, electronic health record, integrated healthcare

Introduction

Health ICT projects are notorious for their failures, despite the efforts of project managers to embed as many critical success factors into the project design as possible. Despite these failures we successfully implement innovations such as the electronic health record (EHR), build on our successes, and sustain their effect.

This paper takes a look at the sustainability of the electronic health record in New Zealand, which is the result of a series of incremental innovative implementations of aspects of the EHR, within the national framework of providing an infrastructure and capacity to build capability for a distributed EHR. The concept sustainability is defined and is linked to change in the complex adaptive system of healthcare. A description is given of the capability crisis that marks the transition period in the change process, and a discussion follows on the relevance of predictions of failure to ultimate sustainable success of an innovation. It is concluded that although we tend to predict failure in the short term, especially during the transition to change, we do not enact our predictions. On the contrary, we adapt to and adopt new processes, technology and relationships that support the growing EHR in order to predispose our healthcare providers to deliver integrated care.

Sustainability defined

There is a general impression that ICT projects tend to fail, with failure rates ranging between 50% and 80%. (1-3) This failure rate appears to apply to health ICT projects as well. There are different descriptions of such failure. A project may be partially successful where it achieves some but not all of its goals or it could simply be an outright failure or be abandoned for any number of reasons. It may be considered a failure because there is no change in an organization after completion of a project, or there may be unpredicted undesirable project outcomes. Some projects are successful in one setting but when implemented in a different setting they fail. Lastly, there are sustainability failures. (3-5)

The term sustainability is used in several contexts and with different meanings. Conversations about sustainability usually include words and phrases such as institutionalization, realization of long-term benefits, long-term continuation of an innovation, financial self-sufficiency, efficiency, and durability. (6) Thus one can assume that an ICT innovation that has been introduced has been sustained if it has been diffused into an organization (7): it has become institutionalized, long-term benefits are occurring, it is financially self-supporting (especially after the funding for the initial project is no longer there) and it is proving to be efficient. It has become part of business as usual, the way we have always done things. It is also essential that the innovation is scaled up to the whole organization and in the case of the electronic health record, the region and whole country wherever appropriate. (8)

Institutionalizing a health ICT innovation is fraught with sustainability issues, especially when funding ends with the completion of an initial project. (8) Political support is essential in the introduction of innovations, particularly long-term support. Once there is executive and leadership support, it is essential that capacity is built and resources are provided for ongoing support and maintenance of the innovation. (6, 8) New technology, processes and relationships need to be institutionalized in order for the innovation to become business as usual.

Sustainability and change

Sustainability is therefore a part of the change process, that is, the change cycle involved in the introduction of an innovation does not end when its implementation project ends. (6) When we change it appears that we follow a process that approximates Lewin's (9) unfreeze, change, refreeze cycle. Most change theories indicate that we follow a process that marks our acknowledgement of the change confronting us and our transition through to final adaptation to the new way of working or living, aiming to achieve a better state than when we started out. (10, 11) The trough in this process, as depicted in Figure 1, can be destructively deep if we attempt to change too much or make a fundamental change without preparatory changes. (4) In order to reduce the impact of such an effect, and to improve the probability of sustained change, we could fill the trough with small accumulative changes so that the larger, more radical changes are laid over, using the preceding ones as their foundation. (4, 12)

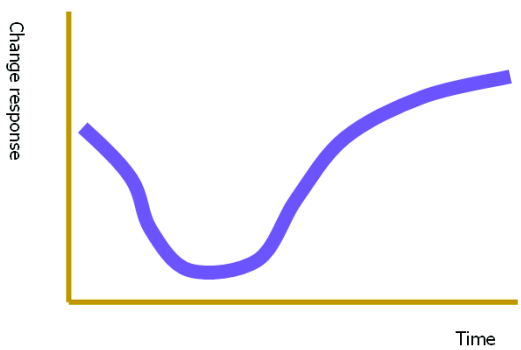


Figure 1 - The change process(10)

In the development of the EHR in New Zealand, the trough of change was filled initially with infrastructure projects that would support future implementation of more components of the EHR. (12) This means that once the ICT infrastructure and foundation administrative IT (information technology) were in place, other more complex clinical applications could be implemented. Incremental changes could be represented in Figure 1 as multiple smaller change curves within the larger curve of the complete EHR. (13) This means that the overall change is cumulative and therefore possibly more likely to be sustained. In this way we accommodate the large change in small increments, thus allowing for the institutionalization of the EHR as we adapt from one change to another, interdependent change. The accumulation of changes results in an accumulation of sustained innovations as each innovation is absorbed as the starting place for the next innovation.

In this way long-term benefits start to appear as we adapt to the changes, as the innovations are absorbed into an organization. (4) The proof of the pudding of sustainability lies in the long-term benefits for the stakeholders and the extent to which they perceive the added value of the innovation being implemented. The challenge for the success of health ICT projects lies more in its sustainability than in the success of the implementation project itself. It lies in the long-term adaptation of the system in which it is introduced rather than in the short-term success of the project. (8) This means that a partially successful project can become fully successful in the long-term, assuming that it achieves the final planned outcomes and sustainability. The accumulative development of the EHR in New Zealand requires a sustained capacity for adaptation to change. In the complex adaptive system of healthcare there is a constant stream of changes that we experience every day. This is compounded by the innovative EHR applications that are being implemented with a view to sustained change in the way we deliver healthcare in a paperless environment. (14)

Complexity and sustainable change

Health can be viewed as a complex adaptive system (15), in which many parts of the system interact interdependently in varying and unpredictable degrees with one another and their environment (16-18). The continuum of complexity ranges from simple and unambiguous with high degrees of perceived certainty, to chaos which extends beyond complexity, uncertainty and ambiguity. Within this context capability is potentially at its best when there is a moderate degree of complexity. (19) It is when things are reasonably complex that change is most stimulating and best received. We usually function well in the position where most of our world is reasonably certain and predictable, fairly unambiguous, familiar, mostly known and knowable, and where interdependencies and relationships are fairly simple. (14) Paradox and tension co-exist in this environment where fuzzy boundaries and sensitivity to initial conditions make it hard to predict the future impact of the implementation of innovations. Emergence of unexpected and unpredictable phenomena is common in health ICT projects, as is characteristic of complex adaptive systems. Changes occur within changes as a result of the butterfly effect (20), which results in disproportionate consequences to even the best laid plans.

The diffusion of a health ICT innovation is not a simple, linear process of plan, implement, review, and reap the benefits. The diffusion itself is fraught with unpredictable consequences, the difficulties of long-term benefit management, and paradoxes that are difficult to assess and exploit. The Health Information Strategy (HIS) for New Zealand (21) attempts to take advantage of these characteristics of healthcare as a complex adaptive system by focusing on the development of a technological infrastructure to support the EHR, parallel to the development and nurturing of technological capability. In this way a distributed EHR is developed that is predisposed to national dispersion and thus the scale required for sustainability. (8) This allows for emergence in such a way that incremental innovations are supportable and scalable according to the capacity of the services in which they are implemented. Once the ICT that is needed for in-house EHR management for one health service is on its way to being institutionalized, an organization is able to make forays

into connecting up with other services, such as primary care organizations. In this way linkages are made between services that allow for more innovations to be implemented. On the surface this approach may appear useful. However, even when attempts are made to maintain manageable projects for implementing these innovations, there are times when the scale of an innovation itself is large, or is perceived to be overwhelming by those involved.

The enormity of change implications

Although we appear to follow a process (depicted in Figure 1) when adapting to our changing and complex environment, the implications of planned change appear to overwhelm most of us during the transition phase of the process. This transition is marked by commitment to planned change but an overwhelming sense that it is all is too much too soon. This transition period usually lasts for a short while (hours or days), but for some people it continues for months. In combination these features appear to comprise a capability crisis. (22)

This capability crisis was researched during the implementation of a large infrastructure project that aimed at establishing a single Information Services platform for an emerging shared services organization for two New Zealand district health boards. (22) Action research was used as a way of managing change in this project in addition to providing a platform for the research to explore the role of the transition aspect of the change process.

The capability crisis described by the project participants is characterized by a heightened sense of ambiguity and complexity where nothing seems certain and everything seems more complicated than ever before – it occurs at the transition period of the change process when people are moving out of their usual way of working and into the complexity of a world in which everything seems to be changing, fluid, unpredictable and difficult. Sensitivity to initial conditions is exhibited in the disproportionate increase in workload, where people are working on multiple levels of the old ways, learning and adopting new ways, and integrating the change into their operational activities. One participant in the project described her experience in an interview towards the end of the project as follows.

It was a terrifying experience because ...it was all very well for other people to be quite glib about that but the scope was creeping. The scope of my task was expanding by the minute.

Paradox and tension are evident in the disproportion between the need and availability of resources. Also, communication appears to be an issue in that people need and demand information, yet they seem to be unable to use it. There is a demand for leadership and yet people appear to be unable to relate to the leadership that is available. People resonate with leaders of their choice rather than taking leadership from appropriate sources as described by another project participant.

He knew he was going to deliver, and he knew it would work and he knew a lot of people thought it wouldn't work. In terms of his leadership there was no question – they could say whatever they like, it was going to work. In terms of leadership and taking a stand that was great.

In addition, predictions of failure mark this crisis: although people are committed to a project's goals and ultimate success, there is a period in which they are uncertain of that success and cannot see how it could be achieved. This tendency to predict failure appears to contradict commitment to project success. It seems that although people appear to be resisting the changes brought about by health ICT innovations, the capability crisis marks their initial attempt at adapting to the change. It is a sign of commitment to the project rather than to the past, the status quo, as described in an interview with a team leader.

You're identifying problems but you don't know there are solutions in the situation. Since you don't know, you feel it's a losing battle.

Predictions of failure

It is during the capability crisis that we realize more fully the implications of our commitment to the project at hand. We suddenly feel inadequate for the task and frustrated that we are unable to learn and develop skills fast enough to be able to make the transition to the new ways of working, using new technology, processes and in new relationships with our colleagues and the project stakeholders. There seems to be a background chatter predicting failure, as described by the training team leader:

There was a lot of talk around me about the project ... "not going to work" ...people around you creating an uncertain environment by the things they say ...making it hard to deal with the situation ....saying things like "how's this going to work? We're never going to get it off the ground."

At the same time as the background chatter of predictions of failure, we hear in counterpoint the management mantra that we must make the project work. Some participants found it hard to relate to what they called "Hitlerism" when the managers did not appear to take their concerns seriously and merely repeated that the project would go ahead and that it would be successful. This was described by an interviewee quoting snatches of conversations she witnessed:

... "but you can't do things like that" ... "but we will do it this way" ... "but you don't understand you can't do it like that because..." "we have to do it, we have a time frame and we have to go ahead" ... "we're rolling..."

Although predictions of failure appear to occur in the early stages of the change process and seem to be resistance to change, we cannot assume that they will disrupt a project or lead to its actual failure. These predictions of failure could be a sign that people are progressing through the early stages of the change process. (10) This means that the commitment to the change brought about by a project remains intact but the predictions of failure may be signaling attempts to accommodate the finer aspects of change in

order to contribute to project success. This is supported by the appearance of frustration and heightened awareness of issues and the inability to perform effectively during a period of change.

Despite protestations that the project would be a failure and that it could not progress to its predicted outcome, people forged on with their project activities. Although there is a temptation to ignore these protestations and predictions of failure, it may be more productive to respond with a mantra that continues to support continued efforts aiming at project success. The capability crisis marks commitment to the project's goals, and therefore its success and eventual sustainability. It is important for managers and leaders to acknowledge the predictions of failure, and to deliberately use these predictions as the background to their forward-looking mantra that focuses project participants on future success. Towards the end of a project the predictions of failure and mantra of success fade as we focus on winding up project activities.

Conclusion

We adopt innovations in many ways, levels and degrees of complexity. The important thing is that the changes brought about by the implementation of the EHR are sustained, especially in a complex system such as healthcare. Sustainability is characterized by the long-term realization of benefits, the adoption of an innovation as part of our daily routine: the institutionalization of the implemented innovation. The adoption of an innovation requires changes to the way in which we work. People appear to follow a change process is which we commit to the change, transition from old ways of working and assimilate the new processes, technology and relationships into our daily activities: the changes become part of how we have always worked.

The transition period in the change process is marked by a capability crisis in which we are overwhelmed by the implications of the project in which we are implementing an innovation. The project marks a foray into complexity that results in a temporary capability crisis. An aspect of the crisis is predictions of failure which is more an expression of adaptation to the innovation than resistance to change. Although there is fear of failure in the early stages of implementation of an innovation, these predictions of failure and the capability crisis are not heralds of project failure – they are more likely a sign that people are making efforts to adapt to the innovation's associated changes. It is up to managers and leaders to keep a look-out for the capability crisis and use it (complete with its predictions of failure) to support efforts to adapt to the changes brought about by such a project. A mantra that encapsulates the project vision should form the counterpoint for these predictions.

What this means for the implementation of the EHR is that large changes could be disruptive and destructive in that they create capability crisis that is too big to handle for individuals as well as for organizations. The incremental approach used in New Zealand has predisposed people working in healthcare to prepare for and adopt innovations in the EHR in an interdependent manner. In this way, successes can be built upon one another, resulting in the building of capability and IT in a sustained manner. As we master one innovation we are able to prepare for and master the next one, consequently absorbing them into our daily work activities. In this way we can exploit paradoxes, tensions and unplanned consequences that emerge as a result of small successes.

As saturation of EHR innovations is achieved within organizations such as district health boards, each service should become capable of making links with other organizations in the interests of integrated healthcare. Small successes accumulate until the distributed EHR develops the capacity to become a nationally connected EHR, thus achieving the scale required for the success of a sustained innovation, the electronic health record.

Acknowledgments

The authors thank project participants at healthAlliance for their contributions to this research.

References

[1] Shore B. Failure rates in global IS projects and the leadership challenge. Journal of Global Information Management. 2005;8(3):1-5.
[2] Glaser J. Why does an IT investment fail to deliver returns? Healthcare Financial Management. 2003;57(4):114 - 21.
[3] Cozijnsen AJ, Vrakking WJ, van Ijzerloo M. Success and failure of 50 innovation projects in Dutch companies. European Journal of Innovation Management. 2000;3(3):150-9.
[4] Heeks R, Mundy D, Salazar A. Why health care information systems succeed or fail. Information systems for public sector management. Working paper series 9. Manchester: Institute for Development Policy Management; 1999.
[5] Keeling R. Project management: An international perspective. London: MacMillan Press Ltd; 2000.
[6] Johnson K, Hays C, Center H, Daley C. Building capacity and sustainable prevention innovations: a sustainability planning model. Evaluation and Program Planning. 2004;27:135 - 49.
[7] Rogers EM. New product adoption and diffusion. Journal of Consumer Research. 1976;2(March):290-301.
[8] Braa J, Monteiro E, Sahay S. Networks of action: sustainable health information systems across developing countries. MIS Quarterly. 2004;28(3):337-62.
[9] Lewin K. Field theory in social science. New York: Harper and Row; 1951.
[10] Elrod II PD, Tippett DD. The "death valley" of change. Journal of Organizational Change Management. 2002;15(3):273-91.
[11] Moore GA. Crossing the chasm: marketing and selling high-tech products to mainstream customers. New York, N.Y.: Harper Business; 1991.
[12] Orr M. Evolution of New Zealand's health knowledge management system. British Journal of Healthcare Computing and Information Management. 2004;21(10): 28-30.
[13] Orr M, Sankaran S. Mutual empathy, ambiguity and the implementation of electronic knowledge management within the complex health system. Systems Thinking and Complexity Science: Insights for Action 11th Annual

ANZSYS Conference: Managing the Complex 2005 5-7 December Christchurch, New Zealand. ; 2005.
[14] Plesk PE, Greenhalgh T. Complexity science. The challenge of complexity in health care. British Medical Journal. 2001;323:625-8.
[15] Dooley K. A complex adaptive systems model of organization change. Nonlinear Dynamics, Psychology, and Life Sciences. 1997;1(1).
[16] Tan J, Wen HJ, Awad N. Health care and services delivery systems as complex adaptive systems. Examining chaos theory in action. Communications of the ACM. 2005;48(5):37-44.
[17] Plesk PE, Wilson T. Complexity science: complexity, leadership, and management in healthcare organisations. British Medical Journal. 2001;323:746-9.
[18] Plesk PE, Greenhalgh T. The challenge of complexity in health care. British Medical Journal. 2001;323:625-8.
[19] Fraser S, Greenhalgh T. Coping with complexity: Educating for capability. British Medical Journal. 2001;323(7216):799-803.
[20] Gleick J. Chaos: making a new science. New York: Penguin Books; 1987.
[21] Health Information Steering Committee (NZ). Health Information Strategy for New Zealand. Wellington: New Zealand Ministry of Health; 2005.
[22] Day K, Norris AC. Supporting information technology across health boards in New Zealand: themes emerging from the development of a shared services organization. Health Informatics Journal. 2006;12(1):13-25.

Address for correspondence

Karen Day,
School of Population Health
University of Auckland
Private Bag 92019
Auckland Mail Centre
Auckland, New Zealand
k.day@auckland.ac.nz

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

A Sustainability View on the EPR System of N. N. Burdenko Neurosurgical Institute

Michael A. Shifrin, Elena E. Kalinina, Eugene D. Kalinin

N. N. Burdenko Neurosurgical Institute, Moscow, Russia

Abstract

This paper aims at the analysis of the "sustainability status" of an electronic patient record system developed at the Medical Informatics Laboratory of the N. N. Burdenko Neurosurgical Institute (EPR/NSI). It includes some of the principles that allowed a small team of developers to create a sustainable EPR system for a large medical institution with complicated diagnostic and treatment processes.

Keywords:

electronic patient record, business processes, database, formalization, sustainability.

Introduction

This paper is concerned with the analysis of factors of sustainability of medical information systems on the example of a single system – the EPR for the N. N. Burdenko Neurosurgical Institute. It takes a shape similar to a clinical case report in which the analysis of a single complicated case helps to understand general rules.

The N. N. Burdenko Neurosurgical Institute is a hospital specializing in the neurosurgical treatment of patients presenting all types of neurosurgical pathology. It has 300 beds in the clinical departments, 40 beds in ICU, 14 operating rooms, diagnostic, radiological and rehabilitation departments, an Outpatient Department, and a full set of diagnostic and laboratory equipment. The Institute carries out a number of important scientific programs in medicine and neuroscience.

The EPR system has been running since 2000. It has about 600 users; it supports the documentation of almost all events, investigations and manipulations taking place during the diagnostic and treatment process (D&TP). The vast majority of paper documents forming the case reports for patients (required under Russian legislation) are hard copies of electronic documents generated in EPR/NSI. There are three main categories of users of the EPR: medical staff (both nurses and doctors), medical administration and scientific researchers. The system is running in 7/24 mode, all the data are inputted at the working places of the users.

Technically the system is based on the Microsoft technologies. It uses MS SQL Server as DBMS, MS IIS as application server, VBScript as programming language and MS Internet Explorer as client. A special piece of technology MEDSET (Modelling, Engineering, Development, Support and Evolution Technology) supporting all stages of the information system lifecycle was developed to guarantee the evolutionary development of the EPR/NSI (see [1 - 3]).

Sustainability features of the EPR/NSI

Sustainability of medical information systems is a complex phenomenon characterized by many features. Below we discuss some features of the EPR/NSI that make this system sustainable.

Lifetime

The EPR/NSI is running for almost 7 years in non-stop mode. It was stopped only once for the weekend for migration to the new system core, but even so all crucial functions of the system were running. Although this timespan is not sufficient to arrive at the ultimate conclusion about the future of the system, it is long enough to confirm the vitality of the principles it was based on.

Evolutionary development

As mentioned above, special efforts were made in order to allow for the evolutionary development of the system. The system was, and is, evolving in three "dimensions":

- "the spatial dimension" : the number of working places increased from 1 at the start (at the Admissions) to more than 400 nowadays, the number of users increased from 4 nurses working in shifts in the Admissions to more than 600;
- "the functional dimension" : the number of functions supporting special kinds of personnel's activity increased from 1 (admission/discharge) to more than 100;
- "the cognitive dimension" – the formalisation of the inputted data changes all the time in order to better espouse the requirements of the different categories of users; it was necessary to achieve the most comfortable proportion of free text fields and fields with a fixed set of values.

Usability of the system (see [4])

The system is actively used by the users: as just mentioned, almost all documents forming the paper case reports are printed copies of electronic documents gener-

ated in the EPR. Moreover, medical administration needs to spend special efforts to have doctors print out all electronic documents because in many cases they are not necessary for everyday activity. The system is actively used by the heads of the departments and the head physician to control the working processes of departments and of the hospital as a whole. The medical statisticians use the option of getting current statistical reports "by click" so that the shifted their attention to more intentional domains (e.g., medical classifications).

The user interface is simple, all Internet users are accustomed to it hence in most cases the teaching efforts are focused on specific features of the workflow organisation. The point is that the migration from conventional paper records to EPR requires more strict data handling discipline.

In fact, seven years of developing the EPR/NSI system in actual practice demonstrated that this system is stable and can resist challenges of reality and can accommodate to permanent changes. All this counts in favour of the sustainability of the system.

The size of the developer team

One more fundamental evidence of the system sustainability is that the team supporting the system has - already for 4 years – consisted of only 4 people: the head of the group (analytical tasks, new users' relations, the testing of new options), a programmer (analytical and programming tasks), and 2 system administrators (users' support, database administrating, testing, as well as teaching at times). This team supports both existing functions of the system and ensures the development and implementation of new ones.

The methodological basis of sustainability

All facts about the EPR/NSI discussed above would be unimaginable without a strong methodological basis. The principal issues of our MEDSET approach are described in this section.

Formalisation

The first principle is that any information system is just a formal representation of a certain part of the real world and human activity in it. Implementation of an information system consists in the immersion of staff's activity in formalized environment. So developers of the system intended to support activity in some subject domain, have to formalize:

- the users' knowledge of processes in the subject domain
- their own knowledge of processes in this subject domain

And, as it always happens in poorly formalised domains like medicine, this knowledge is fuzzy and implicit.

Adequacy of formalisation

Formalisation may be effective only if it is goal-driven. Any complex subject knowledge may be formalised in many ways, accordingly to the formulated goals. Formalisation has to be adequate to the goals.In the case of EPR/NSI, the goal was formulated as follows:

Delivering informational support to everyday doctors' and scientific researchers' activities.

Note that the administrative needs were not included in the list of goals. Our experience demonstrated that requirements of the medical administration may be satisfied almost freely if clinical data are organized and stored in adequate manner.

Ways of knowledge representation

There are only three ways of representing knowledge in an information system: as data structures, as programme code, as stored data. For every information system the proportion of these methods needs to be determined adequately to the goals of the system. The most stable parts of the knowledge may be represented in the database schema, the relatively stable ones in the program code while the most rapidly changing parts may be stored in the database.

Business processes and activity processes

The basic objective for the lot of the EPR/NSI decision was to treat separately business processes (BP) and activity processes (AP). The structure of the first ones reflects the organization of business, the structure of the second ones – the organization of personnel's activities. The BP-structure is relatively stable in a mature institution, and building a model of the data, it is worth being to follow it. The AP-structure is, on the contrary, rather flexible, and it is reasonable to reflect it in the interface model.

In other words, it may be said that the body of BPs reflects the view at the institution from the outside whereas the AP represents the view from the inside. The set and structure of the BP processes is relatively independent on the medical institution: patient registration, multiple laboratory investigations, surgical treatment and so on are typical for a great number of institutions. The APs, by contrast, may be organised in many different ways.

The main results of this separation are listed below:

- the stability of the data base structure and "local" nature of changes required for the implantation of new processes;
- the possibility of the evolutionary development of the system;
- the possibility of a fine tuning of the system interface to the needs of the staff.

The “Drama Principle” for user interface organization

We use the drama principle of "the union of place, time and action" to develop the users' interface, especially its functional part. This means that users' functions are organized in such a way that the screen forms may be filled at one working place in one action. This principle allows us to design simple screen forms containing only those data which are necessary for the user to input the data generated in his activity process.

Conclusion

All of the issues considered above have been essential to the development and implementation EPR system using an extremely small number of developers. We hope that the EPR/NSI will be sustainable until new achievements both in medicine and techniques ultimately force the developers to change their paradigm and to migrate to a new hard-, middle and software.

Now it is necessary to add just one more principle which is the foundation to the success of any information system:

Love the user.

The information system developer should always keep in mind that he works for the sake of the user and not the user works for the sake of his system. It is only under this condition that it is possible to develop a really sustainable system.

Acknowledgments

The authors are grateful to their colleagues, technicians and physicians, whose dedication, patience and good will were crucial for the development and the implementation of the system.

We also thank those who demonstrated criticism to our efforts, for they made us to find better ways of the system development.

Special thanks to Alexandre Vaxman, MS, for help in editing.

References

[1] Shifrin M, Kalinina E, Kalinin E. MEDSET – an integrated technology for modeling, engineering, deployment, support and evolution of information systems. Proceedings of the 15th IEEE Symposium on Computer-Based Medical Systems, Maribor 2002

[2] Shifrin M, Kalinina E, Kalinin E. Electronic patient record for N.N.Burdenko Neurosurgical Institute: on the road to paperless patient record. Proceedings of Medical Informatics Europe – 2002, Budapest, 2003.

[3] Shifrin M, Kalinina E, Kalinin E. EPR project for the NN Burdenko Neurosurgical Institute: goals, technology, results, future trends. British Journal of Healthcare Computing & Information Management 2003; 20(7): 18-21.

[4] Pantazi SV, Kushniruk A, Moehr JR.. The usability axiom of medical information systems. Int J Med Inform 2006; 75(12):829-39.

Address for correspondence

Michael Shifrin
Phone: +7 499 972 8525
Mobile: +7 916 692 6937
Fax: +7 495 250 9350
E-mail: shifrin@nsi.ru

Chapter 9.

Genomics

This page intentionally left blank

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Predicting Coronary Artery Disease with Medical Profile and Gene Polymorphisms Data

Qiongyu Chen*, Guoliang Li*, Tze-Yun Leong*, Chew-Kiat Heng[+]

**Medical Computing Laboratory, School of Computing, National University of Singapore*
3 Science Drive 2, Singapore 117543
[+] Department of Paediatrics, National University of Singapore
5 Lower Kent Ridge Road, Singapore, 119074

Abstract

Coronary artery disease (CAD) is a main cause of death in the world. Finding cost-effective methods to predict CAD is a major challenge in public health. In this paper, we investigate the combined effects of genetic polymorphisms and non-genetic factors on predicting the risk of CAD by applying well known classification methods, such as Bayesian networks, naïve Bayes, support vector machine, k-nearest neighbor, neural networks and decision trees. Our experiments show that all these classifiers are comparable in terms of accuracy, while Bayesian networks have the additional advantage of being able to provide insights into the relationships among the variables. We observe that the learned Bayesian Networks identify many important dependency relationships among genetic variables, which can be verified with domain knowledge. Conforming to current domain understanding, our results indicate that related diseases (e.g., diabetes and hypertension), age and smoking status are the most important factors for CAD prediction, while the genetic polymorphisms entail more complicated influences.

Keywords:

Coronary artery disease, single nucleotide polymorphisms, data mining, machine learning, Bayesian networks.

Introduction

Coronary artery disease (CAD) is the most common form of heart disease in America and Europe [1, 2]. It occurs when the arteries that supply blood to the heart muscles (coronary arteries) become hardened and narrowed. The major risk factors documented for CAD include related diseases (e.g., diabetes and hypertension), family history, gender, plasma lipid, etc [3].

Research in predicting the risk of CAD usually relies on medical profile and family history information. For example, Lapuerta et al [4] used seven different mean lipid values with the neural networks method to predict the occurrence of a complication of CAD. Wilson et al [5] predicted the risk of CAD by identifying risk categories and statistical tests, e.g., linear regression and logistic regression.

With the rapid advancement of biomedical technologies, now we can combine genetic information, such as microarray-based genotyping, to predict the disease risk. This integrative approach gives us better understanding of the fundamentals of the disease. Tham et al [6] combined medical profile, family history and microarray-based genotyping information in a neural network committee approach to predict the risk of CAD, and achieved reasonably good results.

In this work, we studied a set of profiles with comprehensive genetic polymorphism data and non-genetic data, collected with respect to CAD risk analysis. The objective is not only to predict the risk of CAD, but also to infer the dependency relationships among the relevant domain variables. We compared the prediction of several well known classifiers, including Bayesian networks, naïve Bayes, support vector machine (SVM), k-nearest neighbor, neural networks and decision trees. We experimented with predictions on the non-genetic and genetic factors separately, as well as collectively. Two feature selection methods based on Chi-squared test and Gain Ratio show that the non-genetic information, such as diabetes, hypertension, smoking status, and age, are more important for CAD prediction.

To study the dependency relationships among variables, we examined the structure of the learned Bayesian Network. We observed that the learned network identifies important dependency relationships among variables. Most of these relationships can be verified from the information on gene polymorphisms (*e.g.* polymorphisms at different locations in a single gene). Some other dependency relationships deserve further investigation. To our best knowledge, this is the first work to examine the relationships among both genetic and non-genetic variables in the CAD domain with a probabilistic graphical model (Bayesian networks in particular).

Background

Bayesian networks are graphical tools for modeling uncertainty and inferring causal relationships among the variables. The nodes in a Bayesian network represent the

variables in the target domain. The links between nodes represent probabilistic dependencies among different variables, and often imply causal relationships. Together, the graphical representation models qualitative information, while the conditional probability table in each node models quantitative information in the domain.

Bayesian networks can be constructed entirely from domain knowledge, but the process can be very time-consuming. To overcome this problem, there are many research efforts to build Bayesian networks from data [7-9]. For example, to learn the structure of a Bayesian Network, a family of score-based algorithms search through potential network spaces for the network with the best score using some pre-defined scoring function. The scoring function usually reflects how well a network topology fits the data set, e.g., Bayesian Information Criterion (BIC) [10]. Search methods may range from greedy search to exhaustive search, although the latter is not always feasible for Bayesian networks with moderately large number of variables. Other BN structure learning methods include constraint-based learning algorithms that infer network topology based on dependencies among variables. These dependencies are measured by pre-defined statistical tests.

In this study we also considered other major classification algorithms. Naïve Bayes is a probability-based classification method that assumes all the features are independent of each other given the class. A support vector machine (SVM) finds the optimal decision plane by selecting the fewest instances as the support vectors with the largest margin in the feature space. It is one of the most effective classification methods to date, although it sometimes suffers in the presence of noisy data. Neural networks are a type of black-box methods which can approximate any continuous function to arbitrary accuracy provided that the model has sufficiently large number of nodes and the parameters of the model are chosen properly. K-nearest neighbor classifiers are instance-based methods. They make predictions based on the distances between the test data and the training data. The decision tree method is a *de facto* classification method to evaluate other classification methods. It is recursively built based on the information gain of the features.

CAD prediction

The heart disease data set

The data set used in the study is an expanded version of what was originally collected from a Singaporean hospital for predicting CAD using neural networks [6]. It contains information from 2,949 medical profiles. There are 41 variables for each profile, including 10 non-genetic risk factors and 30 candidate gene polymorphisms. The term "non-genetic factors" is loosely used here to include all risk factors that are not directly collected from genotyping. Some non-genetic factors are clearly non-genetic in nature, such as smoking; but some are implicitly related to traits that may be genetic in nature, such as race and family history of disease. Eight of these non-genetic factors are discrete variables and two others are continuous variables. The distributions of continuous variables Age and body mass index (BMI) roughly follow the normal distribution. Both of them are discretized into 9 equal-width categories separately. The non-genetic risk factors are summarized in Table 1.

Table 1- The summary of the non-genetic variables

Code	# of states	Remarks
CAD	2	Healthy or diseased
SEX	2	Male or female
RACE	3	Chinese, Indian and Malay
DM	2	Diabetes, healthy or diseased
HY	2	Hypertension
SM	2	Smoker and non-smoker
FCAD	2	Family history of CAD
FDM	2	Family history of diabetes
FHY	2	Family history of hypertension
AGE	continuous	
CBMI	continuous	body-mass index

Each medical profile also contains 30 candidate gene polymorphisms. Studies have indicated that these gene polymorphisms may affect the patient's chance of developing CAD in Caucasian populations [3]. However, the influences of these genes on Asians (Chinese, Indian and Malay in this study) are unclear. Most of these variations are single nucleotide polymorphisms (SNPs), but there are also a few indels of longer base pairs. In this data set, each gene marker has 3 possible polymorphisms. Some of these markers are different polymorphisms located in various regions of a single gene. We list them in Table 2. Obviously, the included genetic attributes that may affect the chance of suffering from CAD are non-exhaustive.

The class label is CAD, with 0 denoting healthy subjects and 1 denoting patients suffering from CAD. Classification of CAD is based on the presence of at least 50% narrowing in at least one of the major coronary arteries by angiography. Of the 2,949 subjects, 1,462 or 49.6%, are constituted by patients diagnosed with Coronary Artery Disease at the time of data collection; the rest of the subjects were considered healthy at the time of recruitment.

Prediction results

We applied the classification algorithms implemented in the WEKA [11] software package. The entire data set was randomly split into 10 folds for cross-validation. The following results are based on the same 10 folds of data.

Sensitivity is defined as TP/(TP + FN), specificity as TN/(TN + FP), and prediction accuracy as (TP + TN)/(TP + FP + TN + FN), where TP, TN, FP and FN denote the numbers of the true positives, true negatives, false positives, and false negatives, respectively.

Table 2 - Candidate genetic polymorphisms

Code	Gene	Description
G1	Angiotensin Converting Enzyme	Enzyme involved in metabolism of angiotensin
G2	Angiotensinogen receptor	
G3	Angiotensinogen	Precursor of the hormone angiotensin
G4 – G8	Apolipoprotein B (ApoB)	Protein associated with cholesterol
G9 – G10	Lipoprotein Lipase	Enzyme that breaks down fat
G11 – G12	Antithrombin III	Anticoagulating factor
G13 – G18	Fibrinogen	Molecule that forms the blood clot
G19 – G20	Factor VII	Precursor of blood clot formation
G21 – G22	Apolipoprotein A1	Protein associated with cholesterol
G23	Glycoprotein 3A	
G24	5,10-methylenetetrahydrofolate reductase	
G25	Connexin	Protein that composes vertebrate gap junctions
G26	Cholesteryl ester transfer protein	Protein involved in cholesterol homeostasis
G27 – G30	ATP-binding cassette A1 (ABCA1)	Protein involves in cellular lipid removal

Table 3 - Prediction results (Non-genetic variables only)

Algorithm	Sensitivity	Specificity	Accuracy
Bayesian Network	0.886	0.887	0.887
Decision Tree (J48)	0.884	0.884	0.884
K nearest neighbors	0.854	0.896	0.875
Naïve Bayes	0.883	0.886	0.884
SVM	0.896	0.879	0.887
NN (MLP)	0.871	0.853	0.862

To study the effects of non-genetic and genetic factors, we performed three sets of experiments: prediction with non-genetic factors only, with genetic polymorphisms only, and with all the variables. The results are shown in Table 3 ~ 5.

Table 4 - Prediction results (Gene polymorphism only)

Algorithm	Sensitivity	Specificity	Accuracy
Bayesian Network	0.660	0.619	0.640
Decision Tree (J48)	0.625	0.571	0.598
K nearest neighbors	0.603	0.566	0.584
Naïve Bayes	0.672	0.608	0.640
SVM	0.713	0.634	0.653
NN (MLP)	0.611	0.599	0.605

Table 5 - Prediction results (All variables)

Algorithm	Sensitivity	Specificity	Accuracy
Bayesian Network	0.897	0.882	0.889
Decision Tree (J48)	0.884	0.884	0.884
K nearest neighbors	0.772	0.865	0.819
Naïve Bayes	0.889	0.879	0.884
SVM	0.895	0.889	0.892
NN (MLP)	0.877	0.874	0.875

With the same set of variables, the classifiers are generally comparable to one another in terms of prediction accuracy, with SVM showing a slight advantage, and K nearest neighbors being the worst. Further study showed that the resulting receiver operating characteristics (ROC) curves of these classifiers are similar as well. The improvement of these results over those reported by Tham et al [6] may be attributed to the different classification algorithms, and that more data was collected in this study, and thus the patterns in the data are more representative.

The prediction results based on only non-genetic variables are comparable to those based on the entire data set, while the results based on only genetic polymorphisms are considerably worse than the other two. This indicates that in our setting, non-genetic factors have more significant effects on the risk of CAD. There are two possible reasons for this phenomenon. One is that the set of genetic polymorphisms in our data may be a poor representation of the overall genetic factors for CAD, for it does not cover all possible relevant genes. Another reason is that some of the genetic effects were already expressed through the "non-genetic" variables such as race, family history and related diseases. The interactions of genes with other factors, such as smoking and diet habit, are also important. But these interactions are too complex to be analyzed by classification algorithms alone.

Feature selection

To further improve the prediction results, and to identify statistically significant variables, we applied two feature selection algorithms – Chi Squared and Gain Ratio – on the data set. For the most part these two algorithms agree with each other on the ranking of the variables in terms of significance.

Most of the non-genetic variables are highly ranked, except for the family history variables, with FHY (family history of hypertension) being the lowest ranked variable with both algorithms. Domain experts suggested that it may be due to data collection limitation, as the values of these variables come from the subjects' recalled history, instead of medical records. Most of the genetic polymorphism variables are ranked lower than non-genetic variables. Our final selected variable set consists of all non-genetic variables except for FHY, as well as G4, G17, G23, G26 and G30 from genetic polymorphisms variables.

As shown in Table 6, we see a general though slight improvement of prediction accuracy in the new round of classification experiments with feature selection.

Table 6 - Prediction results (with selected features)

Algorithm	Sensitivity	Specificity	Accuracy
Bayesian Network	0.902	0.883	0.892
Decision Tree (J48)	0.884	0.888	0.886
K nearest neighbors	0.847	0.898	0.873
Naïve Bayes	0.901	0.881	0.891
SVM	0.889	0.889	0.889
NN (MLP)	0.869	0.856	0.863

Dependency relationships identified from Bayesian network

Many classification algorithms used in our study are "black box" algorithms. Although they gave reasonably good results, it is difficult to interpret the models built from data. The Bayesian network, on the other hand, includes a graphical representation of the dependency relationships among variables in the problem domain. Figure 1 shows the Bayesian network learned from all the variables using score-based algorithm with hill climbing, without any input of domain knowledge.

The model identifies many important dependency relationships among genetic variables in the study. Some of the polymorphism variables in the group show high dependencies among one another. For example, G13 – G18 are highly correlated to one another and they are closely-connected in the learned Bayesian network. Domain knowledge confirms that G13-G18 are six SNPs in different locations of the same gene Fibrinogen. Another identified group is G27 – G30, which are SNPs in the ABCA1 gene. This group is associated with both "race" and related diseases, which agrees with a prior study on ABCA1 polymorphisms in local population [12].

Our network also identified other patterns, one of which is that some of the gene polymorphisms are irrelevant to CAD when the race of the subject is known in the studied population. For example, G1 (polymorphism in Angiotensin Converting Enzyme) is shown to be blocked from CAD by race. This finding is reasonable since G1 is a silent indel (287 bp, at Intron 16). Therefore, while it is

very likely to be correlated with race, it is unlikely to have an impact on the risk of CAD. Another similar example is the pair G11 and G12, which are both polymorphisms at ATIII. G11 is a 76 bp polymorphism at the 5' un-translated region, while G12 is a silent indel at intron 5.

In addition to the known correlations, the learned Bayesian Network also shows some interesting unknown patterns. For example, there is a strong dependency between G3 (SNP in Angiotensinogen) and G23 (SNP in Glycoprotein 3A). Whether this is a mere artifact or an indication of actual correlation requires further study.

Sensitivity analysis on the final network shows that the top four risk factors for CAD are diabetes, hypertension, smoking status and age, which conform to domain knowledge. No gene polymorphisms are highly ranked in the analysis, which indicates that no single gene makes a major contribution to the risk of CAD. This is certainly a valid finding since CAD is a complex disease with multifactorial etiology and polygenic. No single gene is therefore expected to contribute substantially to the disease risk except in the case of familial hypercholesterolemia.

4. Discussion and future work

The combined effects of genetic and non-genetic factors on the risk of CAD have not been extensively studied in the medical domain. We have applied six well-known classification algorithms for CAD risk prediction in Singapore. Our experiments showed that the prediction results are better than those in a previous study. We have also identified some interesting patterns in the variables. The Bayesian network constructed from structure learning identified some correlated groups among the genetic factors. Most of them can be explained by relevant domain information. In addition, some gene polymorphisms are shown not to have significant correlation with CAD prediction in the studied population based on the data set used. The identified dependency relationships among the genetic factors from the learned Bayesian network call for further studies.

Our experiments showed that non-genetic variables have a more direct influence on CAD, while the effects of genetic factors are more complicated and subtle. This agrees with the general belief that no single gene is a major determinant factor for CAD. The interaction of genetic factors with non-genetic factors, such as smoking, diet and physical activity habit, are important for the disease risk prediction.

The statistical analysis in our experiments has successfully identified several genetic and non-genetic factors that have significant influence respectively on CAD. However, the interplay of these two groups of factors and their combined effects are still unclear. Our study showed that classification algorithms alone may prove to be inadequate in this regard. Bayesian networks, on the other hand, have shown some promising results as effective exploratory data analysis tools. We are also looking at other probabilistic graphical models with richer representation power, e.g., Probabilistic Relational Models [13], that may give us further insights into this problem.

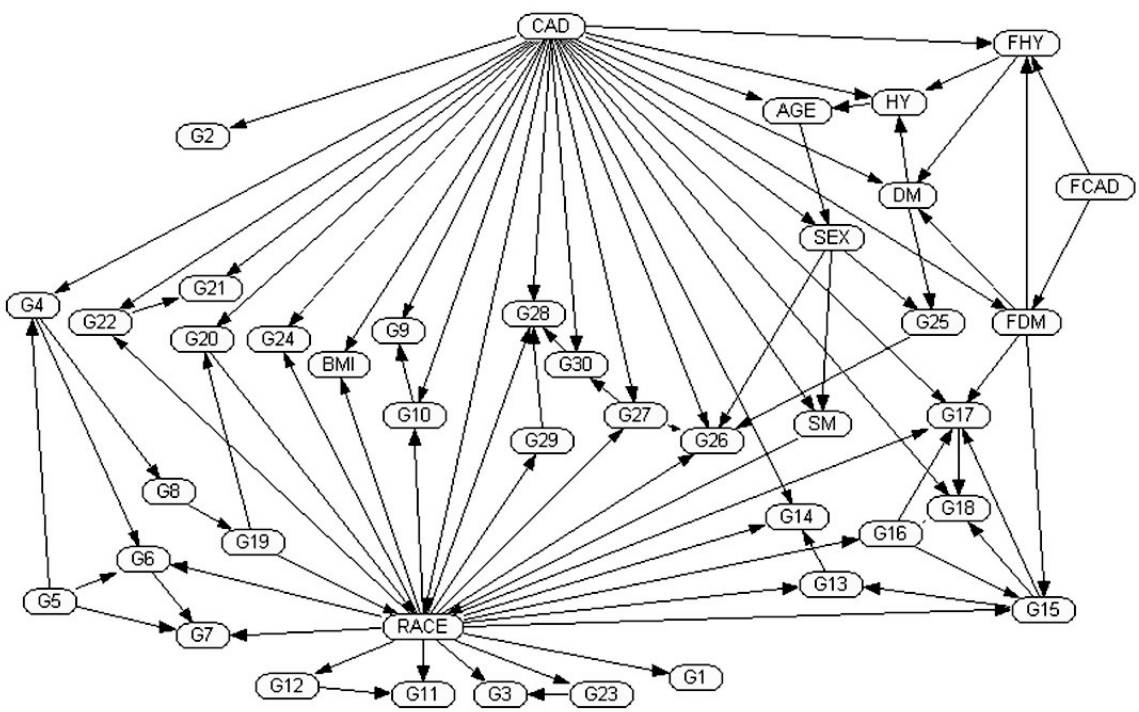


Figure 1 - The learned Bayesian Network

Acknowledgments

We thank Mr. Rohit Joshi for helpful discussions. This research was partially supported by Research Grant No. BM/00/007 from the Biomedical Research Council (BMRC) of the Agency for Science, Technology, and Research (A*Star) and the Ministry of Education in Singapore.

References

[1] Coronary Artery Disease Overview, http://imaginis.com/heart-disease/cad_ov.asp, accessed on February 19, 2005.

[2] Coronary Artery Disease, http://www.nhlbi.nih.gov/health/dci/Diseases/Cad/CAD_All.html, accessed on November 25, 2006.

[3] Heng CK. Candidate Genes for Coronary Artery Diseases, PhD Thesis: National University of Singapore; 1996.

[4] Lapuerta P, Azen SP, LaBree L. Use of neural networks in predicting the risk of coronary artery disease. Computers and Biomedical Research 1995;28(1):38-52.

[5] Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of Coronary Heart Disease Using Risk Factor Categories. Circulation 1998(97):1837-47.

[6] Tham CK, Heng CK, Chin WC. Predicting Risk of Coronary Artery Disease from DNA Microarray-Based Genotyping Using Neural Networks and Other Statistical Analysis Tool. Journal of Bioinformatics and Computational Biology 2003;1(3):521-39.

[7] Cooper GF, Herskovits E. A Bayesian method for the induction of probabilistic networks from data. Machine Learning 1992;9:309-47.

[8] Heckerman D. A Tutorial on Learning with Bayesian Networks. In: Jordan M, editor. Learning in Graphical Models. Cambridge, MA: MIT Press; 1998. pp. 301-354.

[9] Heckerman D, Geiger D, Chickering DM. Learning Bayesian networks: The combination of knowledge and statistical data. Machine Learning 1995;20:197-243.

[10] Schwarz G. Estimating the dimension of a model. The Annals of Statistics 1978;6:461-64.

[11] Witten IH, Frank E. Data mining: practical machine learning tools and techniques with Java implementations. San Francisco: Morgan Kaufmann; 1999.

[12] Tan JH-H, Low P-S, Tan Y-S, Tong M-C, Saha N, Yang H, et al. ABCA1 gene polymorphisms and their associations with coronary artery disease and plasma lipids in males from three ethnic populations in Singapore. Human Genetics 2003;113(2):106-17.

[13] Friedman N, Getoor L, Koller D, Pfeffer A. Learning Probabilistic Relational Models. In: Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI '99): Morgan Kaufmann Publishers; 1999. pp. 1300-1309.

Address for correspondence

Qiongyu Chen
School of Computing, National University of Singapore
3 Science Drive 2, Singapore, 117543
Email: chenqy@comp.nus.edu.sg

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Towards a Top-Domain Ontology for Linking Biomedical Ontologies

Holger Stenzhorn[a,b], Elena Beißwanger[c], Stefan Schulz[a]

[a] *Department of Medical Informatics, Freiburg University Hospital, Germany*
[b] *Institute for Formal Ontology and Medical Information Science (IFOMIS), Saarbrücken, Germany*
[c] *Jena University Language and Information Engineering (JULIE) Lab, Germany*

Abstract

In this paper we present the ongoing development and extension work on BioTop – a top-domain ontology for linking biomedical domain ontologies. We start by making the case for the application of a common ontology to interface independent biomedical domain ontologies by introducing a set of more general classes. Then we briefly depict the relation of BioTop to the GENIA ontology as starting point of its initial developement. Afterwards we propose our distinction of ontologies into top, top-domain and domain ones and describe our approach to the integration of the top ontology BFO into BioTop. Then we present our plans to join the OBO and OBO Foundry repository of ontologies and list its admission principles in relation to our ontology. Some actual BioTop interface classes are shown subsequently. We conclude by detailing on some planned BioTop usages in the area of BioNLP and cancer research and show some further intended improvements.

Keywords:

biomedical ontologies.

Introduction

The last couple of years have seen a tremendous increase in the amount of data collected within the life sciences, especially in biomedicine and its subfield of genomics research. This in turn has spurred many scientific efforts to analyze and structure the newly gained data and to extract further knowledge from it. In the following we are now focusing on the application of ontologies for this particular task and specifically examine one currently existing drawback: Most existing biomedical ontologies – even when having overlapping content – are developed mostly independently from each other. Also, each ontology embraces only some distinct scenario with a mere partial view of the overall scientific field. What has therefore been missing so far is an overarching resource to help with linking and interfacing those independent ontologies. With such a facility in place, new methodologies could be conceived to employ ontologies more efficiently in concert and to create synergetic effects. To this end, we have developed the top-domain ontology BioTop, to be presented in the following.

We concentrated our work on interfacing a smaller selection of about 60 ontologies in the Open Biomedical Ontologies (OBO) framework [1]: the Gene Ontology (GO) [2], the Sequence Ontology (SO), the Cell Ontology (CO), the Chemical Entities of Biological Interest (ChEBI) and the Foundational Model of Anatomy (FMA) [3]. At this point we want to stress that the methodology of our work can be easily applied to create additional interfaces to other ontologies in both this particular topic area as well as similar ones. We are highly interested to further investigate the latter in our future research.

Project background

Relation to GENIA

The initial version of BioTop rested upon the idea to create a comprehensive, formally-based redesign and expansion of the original GENIA ontology [4]. The basic development policy was to follow the fundamental principles of formal rigor, explicitness and precision of ontological axioms and to maintain the clear overall scope of creating a biomedical upper ontology. The implementation was to be based on the Description Logic subtype of the Web Ontology Language (OWL-DL) [5].

The GENIA ontology had originally been developed for and within the biological natural language processing (BioNLP) community and had quickly become a de-facto standard in this field. Its authors claim the ontology to be a formal model of cell signalling reactions in humans and regard its main application to serve as basis for creating thesauri and semantic dictionaries in BioNLP applications (e.g. the semantic annotation of named entities in biological literature abstracts). The GENIA authors also consider it as providing a mutual basis for an integrated view over multiple biological databases.

The GENIA ontology itself is very small, containing only 45 distinct terms, arranged in a simple taxonomy with a maximal depth of six levels. It also limits itself to a set of highly general upper-level classes centred on the notions of biochemical substances and their corresponding locations in the organisms.

During our work, we found some non-trivial shortcomings within the GENIA ontology: Firstly, for most classes a proper documentation and/or textual definition was missing or lacked clarity. Secondly, some class names or their

particular position in the taxonomy contradicted common biological or ontological intuitions of consistency. Both issues can obviously lead to conflicting interpretations and incorrect applications of the classes. A complete analysis of the deficiencies found in GENIA and our proposed solutions can be found in [6].

Top, top-domain and domain ontologies

We propose to distinguish ontologies into three basic types (with their approximate size proportions shown in Figure 1):

- A top ontology (also called top-level or upper ontology) contains only a very small and restricted set of the most high-level, general classes such as "Continuant", "Occurrent", "Function" or "Object" – together with some accompanying relations. Examples for this kind of ontologies are BFO [7] and DOLCE [8].
- A top-domain ontology (also called upper-domain (level) ontology) holds the essential core domain classes to interface to both upper and domain ontologies, like "Organism", "Tissue" or "Cell" in the case of biology. A top-domain ontology can also include more specific relations and further expand or restrict the applicability of relations introduced by the top ontology. An example for this kind of ontologies is BioTop.
- A domain ontology has as its members a multitude of low-level, domain-specific classes to comprehensively describe a certain (aspect of a) domain of interest, e.g. "Antisense RNA Transcription" or "DNA Replication" from the Gene Ontology.

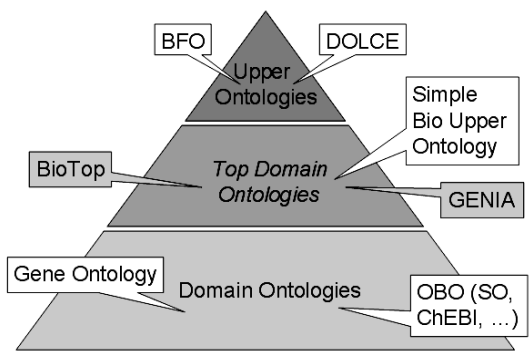


Figure 1 – Ontology layer pyramid (adapted from Alan Rector)

Issues and goals for the second development phase

When we published the initial version of BioTop we received valuable comments that both encouraged us to continue our work on BioTop but also pointed out some existing shortcomings from the standpoint of biomedical and ontology experts. We also received an open invitation to apply for a membership in the OBO Consortium and (subsequently) in the OBO Foundry. Therefore we set up the following list of basic goals we wanted to achieve in the further development:

- Full inclusion and adoption of the Basic Formal Ontology (BFO) and the OBO Relation Ontology (RO) [9].
- Joining the OBO Consortium as well as the OBO Foundry by fulfilling its given set of principles.
- Improving and expanding the existing interface classes to further link the OBO ontologies.

Methods

Integration of top ontologies

Importing BFO and the OBO Relation Ontology (RO)

In the initial version of BioTop we created a set of top-level classes which was based on the terms and definitions found in the publications on both BFO as well as DOLCE (i.e. we initially had a mixture of both top-level ontologies). We also created some additional relations that we based on the ones published in the Relation Ontology (RO).

When the first official version of BFO and RO became available in OWL-DL we decided to modify our ontology by removing the mentioned mixture of BFO and DOLCE top-level classes and to employ the OWL import mechanism to include the now available BFO and RO versions in our ontology.

This integration task was obviously straightforward where actual BFO classes were used before: Here we only had to replace our self-defined class as parent of a given child class with the corresponding class from the BFO import. More problematic were the cases in which DOLCE-inspired classes had been used. Here we had to find either a class in BFO that matched or closely resembled the respective DOLCE class or we had to remodel the class by introducing one or more mediator classes that were deemed to be ontologically sound.

The inclusion of the RO was also straightforward. We had basically only used the relations already defined by the RO and hence simply had to change the references in our ontology from our relation definitions to the imported ones. For the two additional relations we had defined ourselves, we only needed to change the reference to their parent property since they were directly based (i.e. sub-properties) on ones from the RO.

Because of ongoing discussions to consider the addition of new relations to the RO, the need for keeping our self-defined relations might disappear and hence we could remove them. In case the new relations in the RO do not meet our needs we plan to suggest the addition of our relations to the original RO.

Joining the OBO Consortium and the OBO Foundry

After receiving the above mentioned open invitation to join the OBO Consortium we decided to complete the following steps:

- Select the most important published ontologies contained within the OBO (in regard to their user base size and direct relevance to biomedical research), analyze and compare their current top-level classes with the existing interface classes of the initial BioTop version.

Then detect the potential overlap in their respective scope with the scope of BioTop.

- Actively contact the curators and developers of each ontology for which we had developed interface classes and inform them of our plans to join the OBO Consortium and the OBO Foundry and also introduce BioTop as a possible additional layer ontology on top of their ontologies. Additionally invite them to report further ideas on how they think their ontologies and/or BioTop needs to be changed to meet their respective needs.

BioTop and the OBO Consortium and Foundry Principles

In order to apply for a membership in the OBO Consortium, an applicant must take several measures to ascertain that the prospective member ontology fulfils a predetermined set of principles [1]. The following list identifies those principles in relation to the application of BioTop to the OBO Consortium:

1. The BioTop ontology is completely open-source and therefore directly accessible and available to everybody. Links to the latest published version – as well as to older versions – can be found on its website. To adopt BioTop in other projects the respective developer or user must only acknowledge its original source and agree not to alter and distribute the modified ontology under its original name and with the same identifiers.
2. The implementation of BioTop is based on OWL-DL which is now accepted as a common and formally defined language by the OBO Foundry and is established as an official standard for building ontologies for the Semantic Web published by the World Wide Web Consortium (W3C). This in turn entails the availability of a wealth of documentation and supporting tools (for editing as well as classification) and thus allows for a straightforward implementation and adaptation process.
3. BioTop is a top-domain ontology for biomedical research and hence exhibits a clearly defined and delineated subject matter that is distinct from the existing consortium ontologies. Hence it contains only classes and relations necessary to define the higher and (more) general level of this subject field and allows further to link lower level domain ontologies with each other.
4. The string "biotop" is utilized as the unique identifier space for our ontology and serves as the namespace prefix of its OWL-DL implementation. This facilitates avoiding possible naming conflicts as a result of identical class names in BioTop and other ontologies. Also each single class within the ontology holds a unique identifying name to prevent inner-ontology confusion.
5. We include in our ontology precise, plain textual definitions for all classes and relations to resolve the ambiguity that many terms possess in the biomedical sciences. By doing so BioTop cannot only be processed by computer systems but is also understandable for humans and applicable it in their regular work.

In addition to these principles the OBO Foundry requires some additional principles to be followed in order to join:

6. We employ a common version control system to make possible the easy identification and retrieval of all available and different ontology versions. This mechanism simplifies greatly the joint collaboration effort by helping to keep track of all changes happing during the everyday development circle, such as the renaming or the deletion of classes or relations for example.
7. The BioTop ontology uses the relations defined by the OBO Relation Ontology through importing its official, published OWL-DL representation. It additionally introduces two new relations that have been formally defined and strictly follow the prescribed pattern of definitions of the original relations.
8. From the beginning we tried to not solely develop the OWL-DL implementation of the ontology but we were also careful not to overlook the need for a comprehensive and comprehensible documentation. We achieved this in two ways, namely by introducing extensive comments and remarks to each class and relation – directly into the implementation – and also by creating descriptive publications targeted at domain experts with no or little background in ontologies.
9. The interest in our work expressed by various researchers after the first release of BioTop showed us that there exists a multitude of potential users for our ontology. By reaching across several independent domain ontologies all users interested in the combination and interoperability of those ontologies can be regarded as possible users of BioTop also.
10. The development started off as a collaborative effort between researchers from two different institutions. Through face-to-face discussions at conferences, postings on mailing lists and personal mail communication, many more people specialized in ontology and biology are currently getting involved and provide input and comments for the continuous BioTop refinement.

Interfaces to other ontologies

We tried to achieve a comprehensive coverage of the interface classes in BioTop to link together as many biomedical domain ontologies as possible. But we also tried not to lose track of our original goal: We wanted our ontology to be focused on biomedical matters on a more generic level without any limitations to specific subdomains or some particular species.

Thus we wanted to avoid as much overlap as possible with the given domain ontologies wherever possible and sensible from an ontological point of view. When we found a place of considerable overlap we tried to ascertain whether this overlap problem should be solved on the side of BioTop or rather on the side of the domain ontology. So far this problem has not been tackled in a satisfactory fashion but we plan to contact the responsible curators to further discuss matters and to come up with a principled way of handling such cases.

BioTop sometimes does not provide a direct, logical link to the upper-level classes of the domain ontology. In such cases we obviously see the need to talk to the respective ontology curators to find out whether they should intro-

duce some additional top-level classes to their ontology or whether BioTop is missing any important classes that could be general enough to be relevant to other ontologies as well.

Table 1 lists a small sample of the links we have created so far from BioTop to other domain ontologies and which have been accepted as being valid by the respective curators. A more comprehensive treatment of those links can be found in [6].

Table 1 – Example links from BioTop to OBO ontologies

BioTop	**OBO Ontologies**
Biological Process	Biological Process (GO)
Protein Function	Molecular Function (GO)
Cell Component	Cell Component (GO)
Cell	Cell (CO), Cell (FMA)
Atom	Atoms (ChEBI)
Subatomic Particle	Elementary Particles (ChEBI)
Organic Compound	Organic Molecular Entities (ChEBI)
Tissue	Tissue (FMA)
DNA, RNA	DNA, RNA (SO)
Protein	Protein (SO)

Figure 2 depicts some small, restricted portions of the BFO, BioTop and GO ontologies to demonstrate the layering and interfacing in between them. In this particular example the domain-specific GO class "Transcription" is linked to the BioTop class "Biological Function" which in turn is linked to the generic, domain independent class "Function" in BFO.

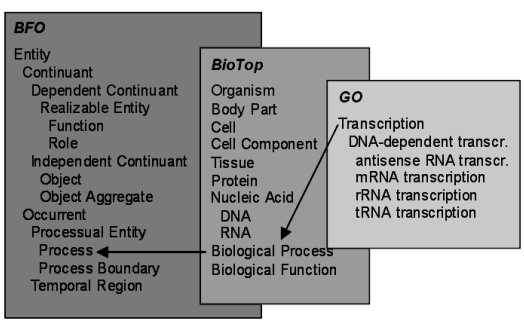


Figure 2 – Classes showing the ontology interfacing

Outlook and Future Plans

Usage Scenarios

For the further evaluation and application of BioTop we are currently pursuing the following two main scenarios.

Natural language processing and named entity annotation

The European project BOOTStrep [11] will apply the BioTop ontology for various natural language processing purposes in the biology domain (BioNLP). Amongst them is its application to improve the quality of semantic annotation of biological text corpora, i.e. the class names from the BioTop ontology are used as the vocabulary to semantically annotate named entities automatically identified in literature abstracts. The resulting annotated corpora are subsequently employed as a training material for statistical methods which in turn build the basis for more complex BioNLP applications such as meaning disambiguation, relation extraction or anaphora resolution.

Involvement in the development of other ontologies

After successfully joining the OBO Consortium and the OBO Foundry we hope that BioTop will become even more publicly visible and subsequently adapted in the development processes of new and existing ontologies in the context of OBO and elsewhere. By doing so, the developers would acknowledge their (and also our) belief in the necessity for interoperability among individual ontologies, as we have stressed above. An application of BioTop by a larger user base will furthermore facilitate the discovery and resolution of residual problems, bugs and shortcomings in the current BioTop implementation.

Further improvements

Amelioration and expansion of interface classes

Future versions of BioTop will incorporate continuously improved interface classes. We are discussing the suitability and applicability of the current interface classes with the curators of several OBO domain ontologies. This will lead to the addition of new or the removal of existing BioTop classes, as well as to the clarification of class definitions or documentations.

One specific interest lies in the present development of an ontology for clinical trials within the OBO Consortium. We believe that BioTop could be applied in this context as a link between biological ontologies on one side and medical ones on the other ameliorating the creation of classes that are both of biological and medical interest. To this end, the first author participates in the creation of an ontology for clinico-genomic trials on nephroblastoma (a specific type of kidney cancer found in children) as part a European research project [12]. This ontology will serve after its completion as major input and the basis for the mentioned OBO clinical trial ontology.

Re-integration of DOLCE

As mentioned above, the initial BioTop version contained a mixture of classes from both BFO and DOLCE at its top level. We now believe that it would be an interesting

experiment to reintroduce DOLCE through including its official OWL-DL version as a second top-layer (in addition to keeping the BFO classes). Firstly, this could perhaps allow us to elicit whether the two ontologies are indeed equivalently applicable as the top-level of our ontology. Doing so could also identify places in BioTop where and why one top ontology might excel the other. Secondly, through the addition of DOLCE we might attract users who are accustomed to this top ontology and are therefore reluctant to use an ontology (solely) based on BFO.

Related work

We are aware of two other projects that are currently engaged in setting up a top-domain ontology for biology. These are the Simple Bio Upper Ontology [13] and GFO-Bio [14]. It seems that at the time of this writing both projects exist in an experimental implementation stadium only and have not produced any publications. Nevertheless we intend to contact both groups for discussions and a possible cooperation.

Our initial intention for BioTop was to improve the interoperability between different biomedical domain ontologies by having a common top-domain ontology. The creation of several top-domain ontologies in this field would obviously be counterproductive and hence some cooperation is essential to achieve a unified solution. Then it would be ideal to have a dedicated workshop to gather the views from more experts in the field to reach a consensus about a single top-domain ontology (which could be based on BioTop or have it as a source).

Conclusion

In this paper we described our current efforts to further develop and extend the biomedical top-domain ontology BioTop. We made the case why an overarching ontology with general classes is important and needed to link independent domain ontologies. We described our integration of BFO into BioTop, discussed our intention to join the OBO Consortium and the OBO Foundry and listed their principles in relation to BioTop. Then we showed some actual interface classes and concluded by detailing on planned BioTop usages and related projects.

Availability

All BioTop material (including its OWL-DL implementation) is available from its website http://www.ifomis.org/biotop.

Acknowledgments

This work was supported by the European Union Network of Excellence "Semantic Interoperability and Data Mining in Biomedicine" (NoE 507505) (http://www.semanticmining.org) as well as the European Commission STREP project "BOOTStrep" (FP6 - 028099) (http://www.bootstrep.eu).

References

[1] OBO Consortium. Open Biological Ontologies (OBO), 2004. [http://obo.sourceforge.net] – Last accessed: March 25, 2007.

[2] Gene Ontology Consortium. Creating the Gene Ontology Resource: Design and Implementation. *Genome Research*, 11(8):1425–1433, 2001.

[3] Rosse C, and Mejino JVL. A Reference Ontology for Biomedical Informatics: the Foundational Model of Anatomy. J Biomed Inform. 36:478-500.

[4] Ohta T, Tateisi Y, and Kim JD. The Genia Corpus: An Annotated Research Abstract Corpus in the Molecular Biology Domain. In Proc. of the 2nd International Conference on Human Language Technology Research (HLT 2002), San Francisco: Morgan Kaufmann, 2002; pp. 82–86.

[5] Horrocks I, Patel-Schneider PF, and van Harmelen F. From SHIQ and RDF to OWL: The Making of a Web Ontology Language. Journal of Web Semantics, 1(1):7–26, 2003.

[6] Schulz S, Beißwanger E, Wermter J, and Hahn U. Towards an Upper-Level Ontology for Molecular Biology. In Proc. of the American Medical Informatics Assn. Annual Conference (AMIA 2006), Washington, 2006; pp. 694–698.

[7] Grenon P, Smith B, and Goldberg L. Biodynamic Ontology: Applying BFO in the Biomedical Domain. In Proceedings of the Workshop on Medical Ontologies, Number 102 in Studies in Health Technology and Informatics, Amsterdam: IOS Press, 2004; pp. 20–38.

[8] Gangemi A, Guarino N, Masolo C, Oltramari A, and Schneider L. Sweetening Ontologies with DOLCE. In Proceedings of the 11th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2002), Siguenza, Spain, 2002; pp.166–181.

[9] Smith B, Ceusters W, Klagges B, Köhler J, Kumar A, Lomax J, Mungall C, Neuhaus F, Rector AL, and Rosse C. Relations in Biomedical Ontologies. *Genome Biology*, 6(5):R46 (1:15), 2005.

[10] OBO Foundry Consortium. Open Biological Ontologies Foundry (OBO Foundry), 2006. [http://www.obofoundry.org] – Last accessed: March 25, 2007.

[11] BOOTStrep Consortium. Bootstrapping Of Ontologies and Terminologies Strategic Research Project (BOOTStrep), 2006. [http://www.bootstrep.eu] – Last accessed: March 25, 2007.

[12] ACGT Consortium. Advancing Clinico-Genomic Trials on Cancer (ACGT), 2006. [http://www.eu-acgt.org] – Last accessed: March 25, 2007.

[13] Rector AL, Stevens A, and Rogers J. Simple Bio Upper Ontology, 2006. [http://www.cs.man.ac.uk/~rector/ontologies/simple-top-bio] – Last accessed: March 25, 2007.

[14] Höhndorf R. GFO-Bio: A Biomedical Core Ontology, 2006. [http://onto.eva.mpg.de/gfo-bio.html] – Last accessed: March 25, 2007.

Address for correspondence

Holger Stenzhorn, Department of Medical Informatics,
Freiburg University Hospital,
Stefan-Meier-Str. 26, 79104 Freiburg,
Germany

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

The Molecular Medicine Informatics Model (MMIM)

Marienne Hibbert [a], Peter Gibbs [b], Terence O'Brien [c], Peter Colman [d], Robert Merriel [e], Naomi Rafael [f], Michael Georgeff [g]

[a] *MMIM Project Director, Melbourne, Australia*
[b] *MMIM Chief Scientist , Melbourne Health and Ludwig Institute, Melbourne, Australia*
[c] *MMIM Neurosciences leader, Melbourne Health and University of Melbourne., Melbourne, Australia*
[d] *MMIM Diabetes leader, Melbourne Health and Walter and Eliza Institute, Melbourne, Australia*
[e] *MMIM Steering Committee Chairman, Melbourne Health, Melbourne, Australia*
[f] *MMIM Senior Database Administrator, Melbourne Health, Melbourne, Australia*
[g] *MMIM Solution Manager, Monash University, Melbourne, Australia*

Abstract

In 2005 a major collaboration in Melbourne, Australia successfully implemented a major medical informatics infrastructure .The convergence of life sciences, healthcare, and information technology is now driving research into the fundamentals of disease causation and toward tailoring individualized treatment. The Molecular Medicine Informatics Model (MMIM) is a 'virtual' research repository of clinical, laboratory and genetic data sets. Integrated data, physically located within independent hospital and research organisations can be searched and queried seamlessly via a federated data integrator. Researchers must gain authorisation to access data, and obtain permission from the data owners before the data can be accessed. The legal and ethical issues surrounding the use of this health data have been addressed so data complies with privacy requirements. The MMIM platform also record links individual cases across multiple institutions and multiple clinical specialties. Significant research outcomes in epilepsy and colorectal cancer have already been enabled by the MMIM research platform. The infrastructure of MMIM enables discovery research to be accessible via the Web with security, intellectual property and privacy addressed.

Keywords[1]:

Medical Informatics, Systems integration, genomics

Introduction

The convergence of life sciences, healthcare, and information technology is revolutionizing the discovery and developmentof new treatments and the optimal use of available therapies [1-2]. Researchers now have the capability to analyse intricate details of human biology through linkage of genetic data with clinical outcomes data giving them the potential to understand the fundamental causation of human disease and predict outcomes. This is driving the development of new drugs, new diagnostics, and lead us to the era of personalised medicine [3]. Key to making the required associations between genotype and phenotype is access to detailed clinical data in sufficient numbers of patients to ensure studies are adequately powered. Few institutions alone have sufficient numbers to perform meaningful analyses, particularly where stratification is performed to look at specific disease attributes. Further, , to utilize all of the available information, clinicians need to look beyond their own specialty into datasets of other disease groupings, analyzing the impact of co-morbidity. To achieve the research objectives promised by this new era in science, sharing of research and clinical data between disparate research groups and institutions becomes critical.

The impetus for the development of MMIM came from a recognition of the need to maximise collaborative research across Australia and internationally. Europe's INFOMED[2], the CaBig in USA [6] and The UK Cancer Grid [7] also recognized the need for collaborative integrated data, data standards and tools. A cohesive approach between disciplines was needed so that research data collection becomes a one time only exercise with the data stored in such a way tha it remained readily accessible, and in a format that facilitated rapid interrogation, permitting multiple research questions across various clinical disciplines and jurisdictions to be addressed. In addition to the value in new and emerging data sets such as genomic data being linked to clinical and outcome data, the is great value to be obtained from existing data. Thus researchers will be able to examine genetic, genomic and proteomic profiles, all factors that may influence treatment outcome, toxicity and potential benefit.

The MMIM Project can enable research from multiple perspectives, including:

- Co morbidities;
- Treatment strategies;
- Genetic predisposition;
- Health Screening activity;
- Genomics, proteomics & epigenetics;
- Outcomes.

1

The objective of the project is to maximise collaborative research efforts, both in Australia and internationally through the development of a federated data integration infrastructure.

Materials and methods

MMIM background

Phase 1 of the MMIM project was successfully completed in 2005. This phase delivered the integration of data across five hospital sites and two medical research institutes involving colorectal cancer, epilepsy, diabetes and the tissue banks.

Phase 2 of the MMIM Project is integrating data across additional Victorian and interstate hospitals including the additional disease types - multiple sclerosis, stroke, cystic fibrosis, asthma, and brain tumours.

Phase 3 of the project is funded over a three year period until June 2009 to provide support for the creation of an Australian Cancer Grid across all tumour types and:

- The Infrastructure expansion – the data grid;
- The research activity and outcomes;

Overview of the MMIM federated mode - technologies

The MMIM project is a federated model where each participating site retains ownership and control over their own data sources and data collection systems. The architecture can be seen diagrammatically in Figure 1.

The data sources used for integration were established as clinical research databases written and maintained by specialist clinicians in their own healthcare facilities. These included datasets that address single questions up to comprehensive datasets capturing information across the full range of a disease, from data regarding genetics and screening through to care of end stage disease. Data was uploaded from these databases nightly, or manually loaded (for static datasets) into a 'cache' database, a local DB2 UDB database termed a Local Research Repository (LRR) located at each site. The distributed LRR databases were federated using IBM Websphere Information Integrator running on a single server termed the Federated Data Integrator (FDI). Information Integrator makes remote databases appear as local DB2 table views, allowing single SQL queries to be executed against all federated data.

Public domain databases were also federated into the system including a local XML flat file and resources from the National Library of Medicine (Genbank, Medline and Uniprot) via an Internet web service. Both data sources appeared relational to the end user even though they were not.

A unique number was assigned to each patient termed a Unique Subject Index (USI) by transferring certain identifying information to Sun (SeeBeyond) e-Ways and replicated to the LRRs via the FDI.

The security system included a number of features. Each LRR was connected to the FDI via Virtual Private Network (VPN) connections, which ensure data privacy and encryption. Views block all identifying information, allowing end users to see only the clinical data in conjunction with the USI. User access to these views on the FDI is controlled by assigning DB2 database roles which define privileges to the table/view level. DB2 Query Patroller is used on the FDI to track all queries for audit purposes. Access is controlled by assigning permission at the table level.

SAS™ Enterprise Guide was used as the interface for researchers to perform queries, statistical analysis and construct reports.

Connectivity:

Each participating research institution has a local data storage facility (the LRR) which are connected via a secure technology involving double encryption, Virtual Private Network and DB2-DB2 encryption. This is the technology commonly used in industries to link various sites of operations for major corporations, such as financial institutions,

Data loading:

On a nightly or ad-hoc basis, the clinical research data is loaded into the LRR at each individual site. This loading process uses an extraction, transform and load software feature that is installed on each LRR.

Record linkage - The Unique Subject Index

The Unique Subject Index (USI) is the key element in linking patient records across disparate data sources within and across institutions. It ensures compliance with privacy. Linking patient/subject records and assigning USI identifiers to data allows patients to be linked across multiple institutions and databases while also observing legal, ethical, privacy and data ownership constraints

The USI was developed based on matching of six key demographic data items (Surname, Given name, Middle name or Initial, Date of Birth, Gender, Digits 5 to 9 of the Medicare Number). The software checks new records for a match against existing subjects, using probabilistic matching and a score is assigned on the basis of match / non-match for each data item. "Fuzzy logic" is used for transpositions, soundex matches, and common "dummy" names (e.g. Babe of, Twin 1). Manual checking of subjects in the "grey area" between thresholds can be undertaken by the data owners.

The USI is a 10 digit unique number assigned by the system to each patient. The process involves sending selected identifying data on a nightly basis from each LRR to a Unique Subject Index (USI) program, where the unique number is generated and stored in an Oracle database. This number is then stored the LRR in encrypted form.

End-user queries:

Researcher access is provided to specified tables in MMIM only on application with the research/purpose fully described and only after approval of the application by the MMIM steering committee and data is only available in de-identified form from the FDI.

The identifying data remains stored on the USI data store with extremely restricted and controlled access. No Health Data is stored externally and only authorized researchers can log in to MMIM via the SAS™ terminal server and send queries to the FDI which 'interrogates' the LRRSs. The results of the query is then put into the researcher's secure folder for statistical analysis (the SAS™ server). All queries are tracked and logged for security audit purposes.

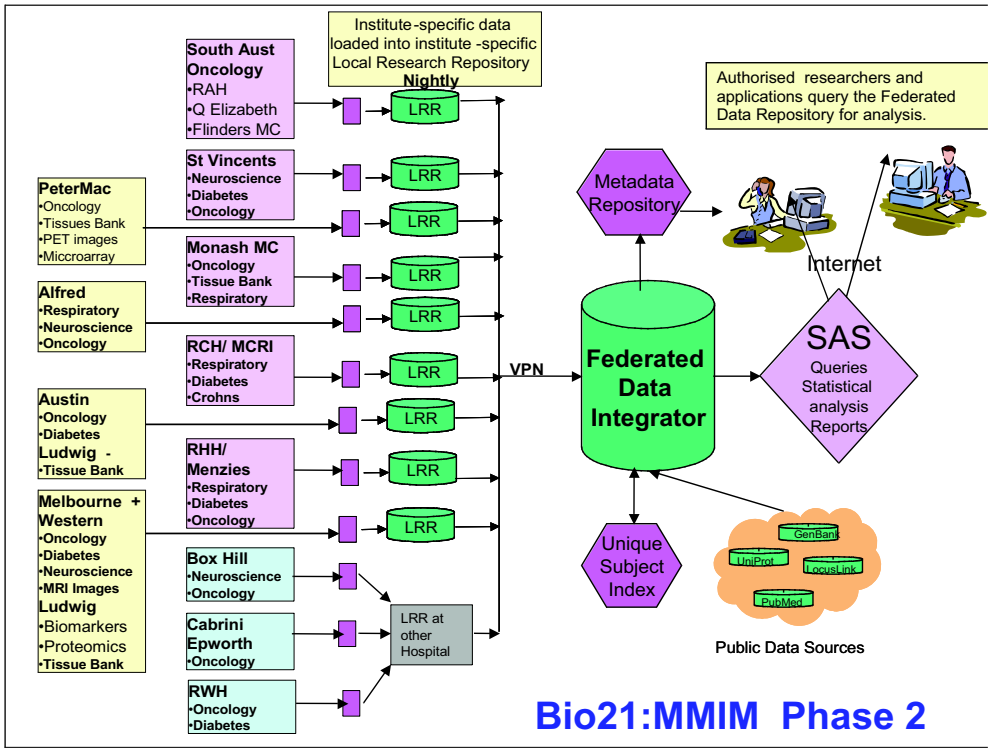


Figure 1. Schematic Architecture of the Bio21:MMIM data grid showing the secure data flow and technologies. Authorized researchers can access the integrated and de-identified data via the Internet.

The MMIM architecture

Diagrammatically the phase 2 architecture is shown in the diagram above.

Security levels:

There are two main levels of security for access to MMIM. Firstly, the users: authorization and usernames and password must be obtained, so only authorised researchers have access to the FDI. Secondly, the MMIM System Administrators who are responsible for building and maintaining the system who have access to all parts of the system (currently there are only two MMIM project authorised personnel with this level of access).

Improving the quality of source data

In MMIM source data was accepted 'as-is' - as provided by participant sites. Standards were used where appropriate (eg NCCI), however provided researchers agreed to the same data definitions, so data was comparable, then specific health data standards were not implemented. However, the databases and systems used for clinical data collection vary at local sites compromising data consistency, but the data quality has improved as the data was available to the clinical researchers. Indeed the project disease/tumour leaders have worked with colleagues in other hospitals and disease areas to standardise data fields and collection processes as far as possible. Further MMIM is working with Australian groups trialing standardised cancer data capture models. These initiatives will contribute to standardised and better quality data over time.

The metadata management – business glossary

The MMIM system provides the ability to search for the range and type of information being captured by MMIM for a particular disease and to discover whether the required data is available - the metadata as opposed to accessing the data itself – the technical data. MMIM users can search and discover the

- clinical areas covered (diseases types& database purpose)
- sites contributing
- types of data collected (pathology, procedures, genetic)
- detail of data elements

and have enough information to request access

MMIM chose IBM's Websphere Business Glossary (WBG) to provide terminology management capabilities. Definitions of standard terms, attaching standard terms to items (databases, tables, or columns), defining the hierarchies of terms, specifying the preferred terms, synonyms and having categories of terms with hierarchies have been implemented. This functionality is important in search and discovery of metadata as users may use non standard or non-familiar terminology (or may misspell words) when describing items. For example, "date of birth" may be described as "dob", "gender" may be described as "sex", etc.

The Business Glossary has open access from the Internet and searching for information can occur by browsing, by drilling down the 'trees' or searching using keywords. In addition the technical - authorized data users have sufficient information about the data items to interpret the data fields, to run queries and to understand the data models so they have the knowledge to join data across databases.

Privacy

The project has continued to seek independent legal advice from lawyers and privacy experts at all stages to ensure that measures taken privacy issues continue to be addressed.

Site and project Human Research Ethics Committee approval has been obtained for all participating sites, and is a prerequisite to proceeding with implementation of MMIM at participating sites..

Intellectual property

A Collaboration agreement that all participating sites must sign to join MMIM explicitly provides for recognition of both Background and Project Intellectual Property.

The project has a set of standard IP management and commercialization processes. Default IP positions are agreed. However, individual research projects are free to negotiate appropriate terms on a case by case basis.

Research results

Examples of research outcomes to date include the areas of epilepsy and neuropsychiatry[8-10], evaluating the sensitivity and specificity of FOBT compared to colonoscopy over a 25 year period [11], and the evaluation of colorectal cancer patients in the areas of biomarkers and therapy [12-15].

Discussion

Building the federated data integration infrastructure

MMIM has successfully built the system infrastructure and federated database integration capability outlined in the methodology section above. This technology allowed the issue of patient privacy, patient record-linkage to be covered as well as ensuring researcher intellectual property was protected. This stage of the project involved data sets (clinical, genomic, tissue bank & biomarkers) for three disease types, namely, colorectal cancer, epilepsy and diabetes.

MMIM is building on this to include a further ten public health sites in Victoria and three states and territories ultimately linking more than 35 disease databases. It is further expanding the technology across the Regional Cancer Services within Victoria and the Metropolitan Melbourne to create the South Eastern Australia part of an Australian Cancer Grid

Powering future research

The MMIM Project has transformed the way that research can be undertaken giving approved and authorised researchers access via the internet to a virtual repository of privacy–protected data not previously available.

From their own work stations researchers can now:

- Link genomic data to clinical / outcome data in Colorectal Cancer and Epilepsy;
- Test multiple hypotheses without collecting / recollecting their own data (with data owner approval);
- Research suitable pre-symptomatic testing and early intervention based on genotype data;
- Research genetic, genomic and proteomic profiles, factors that may influence treatment outcome
- Find out if tissue samples are available for patients with selected clinical profiles
- Analyse summary/statistical information across institutions and from diverse databases.

The following table summarises the traditional approach to research data collection and assembling of databases with that offered by the MMIM Project

Table 1 - Comparative Advantages of Using MMIM

Using Traditional Standalone Research Databases	**Using the MMIM Data Grid**
Static Data at one point in time Often one-off 'dump' Linked once	**Dynamic** Data refreshed & updated Live link to clinical research data Links made on-demand
Often anonymised data	Codified–ethically re-identifiable in exceptional circumstances and privacy protected
Data leaves 'owners' control	Data owners control access
Minimum data sets	Minimum + legacy data
Must specify exact data/ query up-front - can only answer one off specific research question	**Discovery** Research analysis / explorative 'Quality' type clinical reports Clinical Data collected at healthcare site Discovery tools and potential for iterative and exploratory research on a theme (for approved data)
Usually population based studies	Clinical and Biomedical Data collected at healthcare sites and population data

The MMIM virtual repository has enabled collaboration between multiple institutions, both within and across disease specialties, and between clinical researchers, bio-

informaticians and Information Technology specialists. This in turn has expanded research capacity and productivity as follows:

- Within and across disease groups;
- Between data owners;
- Between data owners and researchers across academic institutions in Australia;
- Between data owners and researchers overseas;
- New research data types –e.g. imaging (MRI)

At the same time MMIM can enable expanded teaching and training resources (data sets and tools) in the health research field (medical informatics, genomics, proteomics) which is being developed.

Conclusion

MMIM provides a privacy protected data grid for connecting heterogeneous and dispersed data for medical researchers. The platform is operational and is developing to become scalable and sustainable. It continues to incorporate new data and provide tools for researchers.

Acknowledgments

Phase 1 of t his project was funded by the Victorian Government Department of Innovation, Industry and Regional Development via Bio21 Australia Limited. Phase 2 was funded by the Commonwealth Government of Australian, Department of Education, Science and Training, via their Systemic Infrastructure Initiative. Phase 3 is funded by the Victorian Government Department of Innovation, Industry and Regional Development via the University of Melbourne.

References

[1] Martin-Sanchez F. Iakovidis I. Norager S. Maojo V. de Groen P. Van der Lei J. Jones T. Abraham-Fuchs K. Apweiler R. Babic A. Baud R. Breton V. Cinquin P. Doupi P. Dugas M. Eils R. Engelbrecht R. Ghazal P. Jehenson P. Kulikowski C. Lampe K. De Moor G. Orphanoudakis S. Rossing N. Sarachan B. Sousa A. Spekowius G. Thireos G. Zahlmann G. Zvarova J. Hermosilla I. Vicente FJ. Synergy between Medical Informatics and bioinformatics:Facilitating Genomic medicine for Future Healthcare. Journal of Biomedical Informatics. 37(1):30-42, 2004 Feb.

[2] INFOMED; A virtual laboratory for accessing and integrating genetic and medical information for health applications. EC Project IST-2001-39013. 2002-04 European Commission, Brussels.

[3] Sanz, F, Diaz C, Martin-Sanchez F ,Maojo V. Structuring Eurropean biomedical Informatics to Support individualized Healthcare: current issues and future trends. MedInfo 2004 ;11(pt 2);8030-7.

[4] Maojo V. Kulikowski CA. Reflections on biomedical informatics: from cybernetics to genomic medicine and nanomedicine. Studies in Health Technology & Informatics. 124:19-24, 2006

[5] Maojo V, Garcia-Remesal M, Billhardt H, Alonso-Calvo R, Perez-Rey D, Martin-Sanchez F. Designing New methodologies for Integrating Biomedical Information in Clinical Trials Methods Inf Med 2006;45:180-5.

[6] Von Eschenbach, Buetow K. Cancer Informatics Vision: caBIG ™ Cancer Informatics 2006:2;22-24.

[7] CancerGrid http://www.cancergrid.org/

[8] Yerra R, Kilpatrick C, Matkovic Z, King B, Rafael N, Hibbert M, Brand C, and O'Brien, First Seizure Clinic Experience: Heterogeneity of patient population and prognosis, *Epilepsia* **46** suppl 8 (2005), 360.

[9] Yerra R, Kilpatrick C, Matkovic Z, Belbin S, Rafael N, Hibbert M, Brand C, and O'Brien T, Syndromic diagnosis of epilepsy in the First Seizure Clinic population, Journal of Neurosciences, 238 (2005), S150

[10] Mark Walterfang, Ronald Siu, Dennis Velakoulis. The NUCOG: validity and reliability of a brief cognitive screening tool in neuropsychiatry patients. Australian and New Zealand Journal of Psychiatry 40 (2006), 995-1002

[11] Macrae FA, Slattery MA, Brown GJ, O'Dwyer MA, Murphy C, Hibbert ME, St John DJB. Sensitivity of faecal occult blood testing (FOBT) for asymptomatic colorectal cancer and advanced adenomas over a 25 year experience in colorectal cancer screening. Gastroenterological Society of Australia Australian Gastroenterology Week, Brisbane 2005, Journal of Gastroenterology and Hepatology 20 (2005), A34.

[12] Lim E, Jones IT, Gibbs P, McLaughlan S, Faragher I, Skinner I, Chao MW, and Johns J. Subsite-specific colorectal cancer in diabetic and non-diabetic patients . Cancer Epidemiology, Biomarkers & Prevention 14 (2005), 1579–1582.

[13] Rodrigues J, Lim E, McLaughlin S, Faragher I, Skinner I, Chao M, Croxford M, Chapman M, Johns J, Gibbs P The Influence of Language Spoken on Colorectal Cancer Diagnosis and Management'. ANZ Journal of Surgery 76 (2006), 671.

[14] Gibbs P, McLaughlin S, Skinner I, Jones I, Hayes I, Chapman M, Johns J, Lim L, Faragher Re: Completion of therapy by Medicare patients with stage III colon cancer. . Journal of the National Cancer Institute 98(21) (2006), 1582.

[15] Reiger NA, Barnett FS, Moore JWE, Neo E, Badahdah F, Ryan AJ, Ananda SS, Croxford M, Johns J, Gibbs P The quality of pathology reporting impacts on lymph node yield in colon cancer, Journal Clinical Oncology 25 (2007), 463.

Address for correspondence

Dr Marienne Hibbert, Ph D
MMIM Project Director
6 North, Main Building
Melbourne Health
Grattan Street, Parkville
VIC 3050

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Cancer Genomics Object Model: An Object Model for Multiple Functional Genomics Data for Cancer Research

Yu Rang Park[1], Hye Won Lee[1], Sung Bum Cho[1], Ju Han Kim[1,2*]

[1] *Seoul National University Biomedical Informatics (SNUBI), Seoul National University College of Medicine, Korea*
[2] *Human Genome Research Institute, Seoul University Medicine College of Medicine, Seoul 110-799, Korea*

Abstract

The development of functional genomics including transcriptomics, proteomics and metabolomics allow us to monitor a large number of key cellular pathways simultaneously. Several technology-specific data models have been introduced for the representation of functional genomics experimental data, including the MicroArray Gene Expression-Object Model (MAGE-OM), the Proteomics Experiment Data Repository (PEDRo), and the Tissue MicroArray-Object Model (TMA-OM). Despite the increasing number of cancer studies using multiple functional genomics technologies, there is still no integrated data model for multiple functional genomics experimental and clinical data. We propose an object-oriented data model for cancer genomics research, Cancer Genomics Object Model (CaGe-OM). We reference four data models: Functional Genomic-Object Model, MAGE-OM, TMA-OM and PEDRo. The clinical and histopathological information models are created by analyzing cancer management workflow and referencing the College of American Pathology Cancer Protocols and National Cancer Institute Common Data Elements. The CaGe-OM provides a comprehensive data model for integrated storage and analysis of clinical and multiple functional genomics data.

Keywords:

cancer, genomics, data model, standards

Introduction

Functional genomics includes studies of the abundance of gene transcripts by microarrays (transcriptomics), the abundance, localization and interactions of the translated proteins (proteomics), the flux in related metabolites (metabolomics), and various others [1]. For managing and representing of these functional genomics data, several technology-specific data models have been proposed, including MAGE-OM for microarray [2], PEDRo for proteomics [3], SMAR [4], ArMet [5] and MIAMET [6] for metabolomics, and TMA-OM [7] for tissue microarray.

Current researches emphasize the need to integrate data from multiple functional genomics [8, 9]. Following these trends, many cancer researches have been conducted using multiple functional genomics technologies including DNA microarray, 2DE/MS and Tissue Microarray for the understanding of global biological characteristics [3-5].

As the number of cancer studies using multiple functional genomics technologies increases, there are increasing demands for flexible solutions for systematic management of these data. Several databases have been implemented for specific functional technologies or specific purposes [12]. Despite the necessity, there is a few number of integrated data models (Table 1). It only supports a few genomics technologies or document models for genomics and clinical data. Furthermore, most approaches are designed without consideration of extendibility for integration with new biological data models.

In the present study, we designed a new data model for cancer genomics research using multiple functional genomics data, Cancer Genomics Object Model (CaGe-OM).

Table 1 - Previous approaches for integrated model

	Method	Target data	Reference model	Implementation
FGE-OM (Jones A et al., 2004)	Integrated object model	Transcriptomics, and proteomics (2DE and MS)	MAGE-OM, PEDro, Gla-PSI	RAPAD (microarray, 2DE and MS data)
SysBio-OM (Xirasagar S, et al., 2004)	Integrated object model	Transcriptomics, proteomics and metabolomics	MAGE-OM, PEDro, and a model for protein-protein interaction and metabolite	CEBS (only for microarray data)
Genotype Shared Model (HL7 clinical genomics SIG)	Document (XML)	Transcriptiomic, proteomics, sequence and clinical data	HL7 CDA	Genetic testing : BRCA Tissue typing: BMT
IBM GMS (Robson B, et al., 2004)	Document (XML)	Clinical and genomics (protein structure) data	HL7 CDA	Genomic Messing System Language (GMSL)
caCORE (Covitz PA, et al., 2003)	Object oriented API (caBIO)	Clinical and genomics data	Object Model	caBIG, CGAP, MMHCC, caArray etc..
XDesc (Shifman MA, et al., 2004)	EAV and Relational model	Clinical and genomics (Transcriptomis) data	TrialDB	YMD

Materials and methods

We used class diagram of Unified Modeling Language (UML) to represent the concepts, objects and relationships in multiple functional genomics data for cancer research. We reference four experimental data models (FuGE-OM, MAGE-OM, PEDRo and TMA-OM) and two clinical and histopathological data models (College of American Pathologist Cancer Protocols and National Cancer Institute Common Data Element) to design a data model for cancer genomics research.

Functional genomics experiment data modeling

For designing a framework to represent results from multiple functional genomics investigation, we reference four data models; the FuGE-OM for common aspects of experiments, the MAGE-OM for microarray, the PEDRo for proteomics, and the TMA-OM for tissue microarray. The FuGE-OM focused on modeling the common artifacts of functional genomics, such as sample preparation, protocols, instruments, and contact details [1]. Following the wisdom of the FuGE-OM, we reference packages associated with common aspects of functional genomics in three data models (MAGE-OM, PEDRo and TMA-OM) and modify the existing packages and classes within FuGE-OM for describing common biological experimental data which belongs to CommonBioData namespace. We extract technology-specific packages for each three data model. These extracted packages comprised in TechnologySpecific namespace.

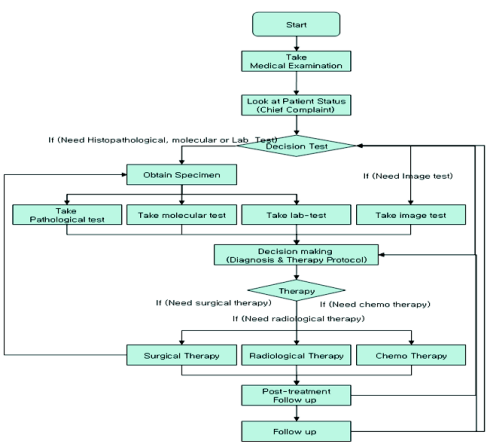


Figure 1 - Workflow diagram of clinical management of cancer. Diamonds indicate events and rectangles are physical entities.

Clinical and histopathological data modeling

For designing the clinical and histopathological data modeling, we analyzed cancer management and referenced document models of clinical and histopathological information, such as College of American Pathologists (CAP) Cancer Protocols (CPs) [13] and National Cancer Institute (NCI) Common Data Element (CDE)[14]. Figure 1 shows the workflow diagram of clinical management for cancer. To obtain comprehensive and extensible data models, we have created 6 packages (i.e., MedicalExamination, Histopathol, Specimen, DecisionTest, Therapy, and FollowUp) by systematically capturing the event and process from the workflow diagram (Figure 1) and category and value from the 43 CAP CPs and NCI CDE.

Results

Workflow of clinical management of cancer

For structured modeling of clinical data for cancer, it is required to analyze workflow of clinical management for cancer like diagnosis and therapy (Figure 1). When a patient arrives at a hospital, she/he takes a medical examination (captured by MedicalExamination class). Medical examination is an event to look at a patients status by a doctor based on physical examination (i.e. inspection, auscultation, and palpation) (PhysicalExam). Then the doctor writes down chief complaints of the patient. Then the patient takes a decision test such as clinical laboratory, images, histopathological and molecular tests (DecisionTest). The doctor makes a decision about the diagnosis and therapy protocols based on the results from various decision tests (Diagnosis & Plan). There are three types of therapy: surgical, radiological and chemotherapy (Therapy). In solid tumor, tumor specimen is obtained after surgical therapy. And then, histopathological test is taken on specimen (HistoPathol). After the therapy, post-treatment follow up is taken on the patient (FollowUp). After primary treatment, the patient is observed according to the follow up schedule (FollowUp).

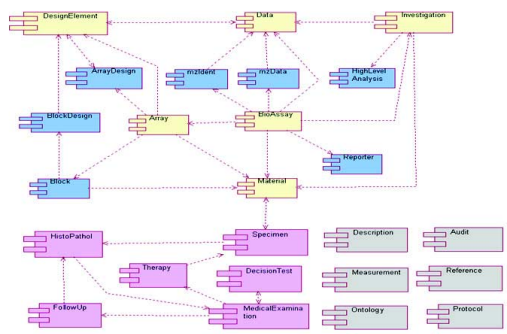


Figure 2 The relationships of 26 packages in Cancer Geno-mics Object Model (CaGe-OM). Most packages in this model are catego-rized into three namespaces; the CommonBioData (yellow-colored), ClinicalData (pink-colored) and Technolo-gySpecificData (blue-colored) namespae. Six packages (gray-colored) are remaining for general purpose.

Overview of cancer genomics object model

CaGe-OM is a data model that contains 183 classes grouped into 25 packages. Figure 2 shows the relationship of CaGe-OM packages, which grouped in three namespaces: the CommonBioData, TechnologySpecific, and ClinicalData models. The CommonBioData contains a

set of packages that describe common aspects of functional genomics from microarrays, proteomics, tissue microarray or potentially other functional genomics techniques. The packages belong to ClinicalData namespace represents clinical and histopathology data of cancer. The TechnologySpecific namespace contains the packages for describing technology-specific data. The remaining 6 packages for common annotation: Description, Measurement, Ontology, Audit, Reference, and Protocol. This model has three abstract classes at the top-level, Extendable, Describable and Identifiable that are unchanged from MAGE-OM. Most classes inherit their attributes. CommonBioData and TechnologySpecificData can refer to ClinicalData through Material packages in CommonBioData namespace.

CommonBioData namespace

The CommonBioData namespace represents common aspects of functional genomics experiments. Experimental design (captured by Investigation package), biological sample preparation (Material) and biological molecules such as DNA or protein sequences (ConceptualMolecule) are common components of all functional genomics investigation. The CommonBioData namespace consists of six packages; Investigation, Material, Array, DesignElement, Data and BioAssay.

To represent data from experiments using any type of technology, packages contained in this namespace have a generic structure. In Data packages, for instance, the Data object represents a container for a set of multidimensional data matrices and the coordinate set found in each of the dimensions.

The Material is a package for all biological and physical materials involved in an experimental workflow. For integrating genomics and clinical data, this package has a relationship with Specimen package that belongs to ClinicalData namespace. As a result, CommonBioData namespace has a reference to ClinicalData via Material and Specimen classes, representing the clinical and histopathological information of the specimen used in functional genomics experiments.

The Array, ArrayDesign, and DesignElement in the MAGE-OM contain information regarding the design, manufacturer and contents of microarrays (http://www.omg.org/docs/formal/ 03-02-03.pdf). In these packages, the ArrayDesign package is a microarray specific package. However, Array and DesignElement is commonly used in the functional genomics experiments such as microarray, tissue microarray and proteomics. Thus we are adding these two packages into CommonBioData namespace.

Clinical Data namespace

The Clinical Data namespace includes package with classes covering clinical and histopathological data 43 cancer types considered by CAP CPs, and clinical contexts from the workflow analysis of cancer management and NCI CDEs. The ClinicalData namespace is composed of six packages; MedicalExamination, Histopatho, Specimen, DecisonTest, Therapy and FollowUp.

The MedicalExamination package, the core package in ClinicalData namespace, contains classes for Demography, PhysicalExam, History, Diagnosis and Plan. Through MedicalExamination classe, all the other packages contained in ClinicalData namespace have associations with Medical-Examination package (Figure 3(a)).

The HistoPathol package provides classes describing histopathological information of specimens (Figure 3(b)). The BasicHistoPathol class stores elements that should be included regardless of the organ and tissue. The OrganSpecific class store elements for specific organs. The BasicHistoPathol class is an abstract class, the subclasses of which are the TumorInfo and Histology classes.

Classes in DecisonTest package are describing several medical tests such as image, laboratory, molecular and histopathological test. Therapy package contain classes to store data from medical therapy; surgical, radiological, and chemotherapy. Specimen package provide classes describing information of tissue obtained by surgical therapy or biopsy. The FollowUp package defines the classes for follow up data like recurrence and vital sign of patient.

TechnologySpecificData namespace

For storing technology-specific data of the experiment, the TechnologySpecificData namespace contains the packages derived from MAGE-OM, PEDRo, and TMA-OM. The TechnologySpecificData namespace is composed of eight packages; ArrayDesign, HighLevelAnalysis, Assay, mzData, mzIdent, Block, BlockDesign and Reporter.

The ArrayDesign and HighLevelAnalysis packages are microarray-specific packages. These packages are reused from corresponding MAGE-OM packages. ArrayDesign includes the manufacturing protocols, contacts, and details of the exact materials used for each feature in Array. The HighLevel-Analysis is a package for the analysis results.

The Assay, mzData and mzIdent packages, which come from the PEDRo, are proteomics-specific packages. The Assay package provides classes and attributes that contain information and annotation on the event of proteomics experiment using 2D or MS and the acquisition of images. The mzData package stores the output from mass spectrometry (MS). The mzIdent package contains the output (and input parameters) from database searches with mass spectrometry data to identify proteins or to quantify protein abundance.

The Block, BlockDesign and Reporter are tissue microarray-specific packages. These packages are identical to packages of the same name in TMA-OM. The BlockDesign package stores the intended pattern of individual block elements. The block with large number of tissues is constructed according to the BlockDesign and the block is sliced to arrays. The Block package records information on the actual events manufacturing blocks. The Reporter package contains classes for reporters used in TMA experiments. The reporter represents materials to identify a particular molecule like gene, protein, or DNA sequence.

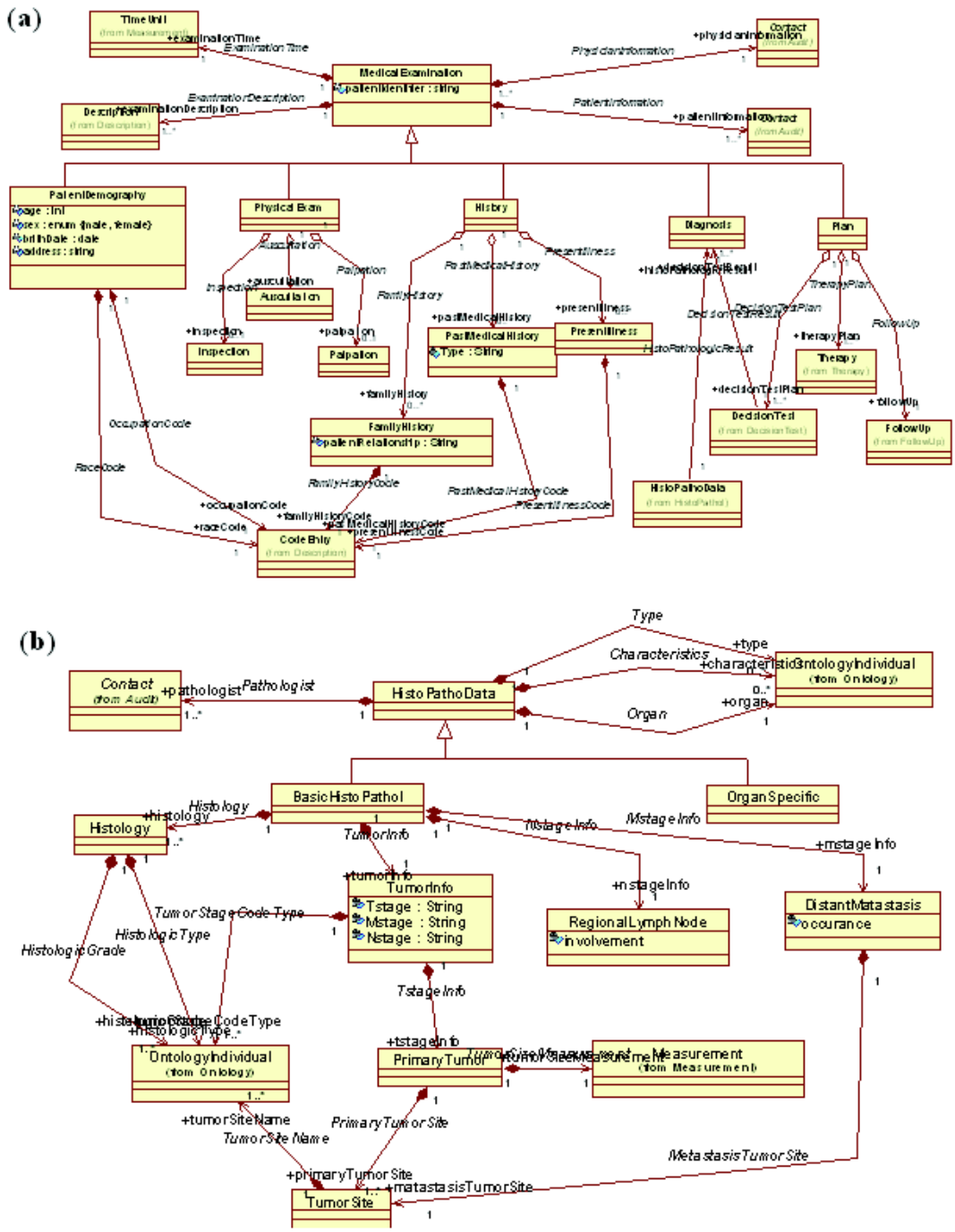


Figure 3 - Class diagram of (a) MedicalExamination and (b) Histopathol packages in ClinicalData namespace.

Discussion

We developed a data model, CaGe-OM, to store and integrate data generated from microarray, proteomics and tissue microarray experiments performed on the same biological samples. The CaGe-OM can represent clinical and histopathological information as multiple functional genomics data for any type of cancer. This integrated data model allows the combined analysis of multiple functional genomics data for understanding of the underlying biological nature in a systematic fashion.

The CaGe-OM can integrate easily a new biological data model without significant difficulty by representing common aspects of the new models as CommonBioData and technology-specific parts as TechnologySpecificData separately, while it is hard to modify the models in previous studies to condider and integrate a new model (Table 1). Because the CaGe-OM is independent of implementation, several applications based on this model such as relational database schema, web application and XML document can be constructed.

The development of an integrated data model for cancer genomics researches may facilitate tight integration of technology-specific data models and clinical data models. As functional genomics are increasingly used in cancer research, the CaGe-OM will be useful for the structured data management of clinical data and for the analysis of functional genomics data combined with clinical data.

Acknowledgments

This study was supported by a grant from Korea Health 21 R&D Project, Ministry of Health & Welfare, Republic of Korea (0412-MI01-0416-0002).

References

[1] Jones AR, Pizarro A, Spellman P, et al. FuGE: Functional Genomics Experiment Object Model. OMICS. Jun 2006; 10(2): 179-84

[2] Spellman PT, Miller M, Stewart J, et al. Design and implementation of microarray gene expression markup language (MAGE-ML). Genome Biol. Aug 2002; 3(9):Research0046.1-9

[3] Garwood K, McLaughlin T, Garwood C, et al. PEDRo: a database for storing, searching and disseminating experimental proteomics data. BMC Genomics. Sep 17 2004;5(1):68.

[4] Castle AL, Fiehn O, Kaddurah-Daouk R, et al. Metabolomics Standard Workshop and the development of international standards for reporting metabolomics experimental results. Brief Bioinform. Jun 2006;7(2):159-65.

[5] Jenkins H, Hardy N, Beckmann M, et al. A proposed framework for the description of plant metabolomics experiments and their results. Nat Biotechnol. Dec 2006;22(12):1601-6.

[6] Bino RJ, Hall RD, Fiehn O, et al. Potential of metabolomics as a functional genomics tool. Trends Plat Sci. Sep 2004;9(9):418-25.

[7] Lee HW, Park YR, Sim J, et al. The tissue microarray object model: a data model for storage, analysis and exchange of tissue microarray experimental data. Arch Pathol Lab Med. Jan 25 2006; 130(7):1004-13

[8] Ideker T, Galitski T, Hood L. A new approach to decoding life: systems biology. Annu Rev Genomics Hum Genet. 2001;2:343-72.

[9] Waters M, Boorman G, Bushel P, et al. Systems toxicology and the Chemical Effects in Biological Systems (CEBS) knowledge base. EHP Toxicogenomics. Jan 2003;111(1T):15-28.

[10] Callagy G, Pharoah P, Chin SF, et al. Identification and validation of prognostic markers in breast cancer with the complementary use of array-CGH and tissue microarrays. J Pathol. Feb 2005;205(3):388-396.

[11] Ippolito JE, Xu J, Jain S, et al. An integrated functional genomics and metabolomics approach for defining poor prognosis in human neuroendocrine cancers. Proc Natl Acad Sci U S A. Jul 12 2005;102(28):9901-6.

[12] Elfilali A, Lair S, Verbeke C, et al. ITTACA : a new database for integrated tumor transcriptome array and clinical data analysis. Nucleic Acids Res. Jan 1 2006;34(Database issue):D613-6.

[13] Warzel DB, Andonaydis C, McCurry B, Chilukuri R, Ishmukhamedov S, Covitz P. Common data element (CDE) management and deployment in clinical trials. AMIA Annu Symp Proc. 2003:1048.

[14] The College of American Pathologists. Cancer protocols in January 2006 revision version. Available at: http://www.cap.org/apps/docs/cancerprotocols/protocolsindex.html. Accessed November 8, 2006.

Address for correspondence

Ju Han Kim, MD, PhD,
Seoul National University College of Medicine,
28 Yongon-dong, Chongno-gu, Seoul 110-799, Korea.
Tel: +82-2-740-8320; Fax:+82-2-742-5947
e-mail: juhan@snu.ac.kr

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Automatic Pedigree Reconstruction for Genetic Studies in Isolated Populations

C. Larizza[a], I. Buetti[b], G. Milani[a], A. Nuzzo[a], C. Sala[b], D. Toniolo[b], R. Bellazzi[a]

[a] *Department of Computer Science and Systems, University of Pavia, Pavia, Italy*
[b] *DIBIT - San Raffaele Scientific Institute, Milan, Italy*

Abstract

This paper describes a tool implemented to automatically reconstruct the pedigree of an isolated population of Northern Italy with the aim of supporting genetic studies. The goal of such studies is to analyze genealogic, clinical and genetic data for genetic dissection of complex diseases. In this context the reconstruction of the population pedigree is fundamental to verify that such population is a genetic isolate and obtain the parental relationships among the individuals participating to the study. The algorithm presented in the paper, from heterogeneous data sources (demographic municipal and parish archives and other data sources), derives the pedigree applying several heuristic rules in a predefined order. One of the main difficulties in performing such task stands in the "record linkage" process that requires the definition of a sufficiently general strategy for managing the ambiguities caused by missing or imprecise/erroneous input data. The paper, finally, presents and discusses the preliminary results obtained by reconstructing the pedigree of four villages from the data collected during the first eighteen months of project.

Keywords:

pedigree, record linkage, isolated population.

Introduction

A specific interest of the post-genomic era is the correlation of genotypical and phenotypical information with the aim of investigating the genetic components of complex disorders. This research activity requires a strong involvement of Biomedical Informatics and, more in general, of Informatics to provide knowledge and tools for dealing with such an ambitious goal [1].

The term complex disorder refers to the phenotypes resulting from abnormalities that cannot be considered as dependent from a single gene, as in Mendelian diseases [2]. In this case, the problem of understanding their nature is typically very complex, since the risk associated with a mutation may depend for a large part on interaction with other genetic or environmental risk factors. As a consequence, it is usually impossible to find a genetic marker which turns out to be a perfect predictor of the complex trait. Different approaches can be applied to cope with this problem, ranging from linkage analysis to association studies [3]. Those studies, however, may be hampered by several factors, in particular by the small effects of the genetic variants on complex disorders and by the genetic heterogeneity of such variants [4].

In recent years there has been growing interest in mapping disease genes in genetically isolated populations, in which, due the small number of founders, the genetic heterogeneity is reduced, the environmental noise is usually minimized, and a wide set of genealogical data may be available [5]. For this reasons, many concurrent studies on isolated populations have nowadays started in Europe [6-7], Canada [8] and United States [9]. Within these studies, a cohort of citizens (or all citizens) belonging to the isolated population is visited, screened and genotyped. The information about their family trees is derived from public registries government and from other sources, such as the parishes' archives. The final goal of such studies is the joint analysis of genealogical, clinical and genetic data with the aim of obtaining long term research outcomes and, at the same time, short term public health outcomes for the population under screening. In such studies it is evident that the correct reconstruction of the population genealogic tree is a fundamental step for two different reasons: first, because it allows verifying whether the population has few founders and, therefore, can be considered genetically homogeneous, as expected; second, because it allows finding parental linking among individuals belonging to the same phenotype and therefore identifying their common ancestors.

Given the great number of individuals usually involved in such studies (usually the number ranges from 2.000 to 3.000 living people) and the need of reconstructing their pedigree at least from 1600 (the total number of individuals to be considered can increase up to 100.000), it is very important to be able to perform such task automatically or by minimizing, as much as possible, the manual intervention. However, the intrinsic characteristics of such kind of populations (e.g. a little set of frequently recurrent names and surnames) and of the available data sources (e.g. paper registries containing often imprecise or illegible data, very frequent name/surname variations due to transcription errors, language evolution or poor quality of the archives) require the adoption of specific strategies. In particular, it is necessary to choose a correct and robust approach for performing "record linkage" [10], basic processing step of the pedigree reconstruction. Record linkage is the task of quickly and accurately identifying records corresponding to the same entity (in our case, the entities are individuals) from one or more data sources, like municipal and parishes archives.

In this paper we describe the algorithm for the pedigree reconstruction, defined for an Italian project started in 2005. It involves the DIBIT-Hospital San Raffaele (HSR) of Milan and the Laboratory for BioMedical Informatics of the University of Pavia in studying the genetic component of complex disorders in the population of the Val Borbera, an isolated valley in the North of Italy. The aim of the paper is to present the different problems encountered in deriving the pedigree of an isolated Italian population and to discuss the computational solution adopted within the project. Moreover, the first results obtained by reconstructing the pedigree of the population since 1838 and an evaluation of the performance of the algorithm are presented. Finally, the open problems and the future developments aimed at improving the algorithm are discussed.

Materials and methods

The Val Borbera project

The Val Borbera is an isolated valley in the Appenine Region in Northern Italy that has been geographically isolated from the surrounding area until recent years. The population was mainly farming until 50 years ago and seems to have had a constant, but limited, growth over the centuries until a large portion immigrated to the Americas at the beginning of the last century. The part of the population that did not emigrate is still living in the valley or in the surrounding and corresponds to about 2000 people. We started in June 2005 collecting the acts of the municipal archives of four of the eight little towns and the birth and marriage certificates of the several little parishes of the valley, by transferring them from paper to electronic form. Moreover, we exploited the electronic demographic archives available since 1985 to manage the visits booking of the people participating to the study. In the future we will record also the death certificates of the population, which data will be exploited to refine and improve the reconstruction algorithm.

Computational methods

This section describes the different data sources available for the pedigree reconstruction, the architecture of the database designed to support the different phases of the algorithm, the problems encountered during the analysis of the data and the computational solution adopted.

The data sources

The study exploits the demographic information coming from municipal and church archives, quite easily traced in the few towns and parishes of the valley, stored in an MS-Access database. Such data represent the input to the pedigree reconstruction algorithm. The overall data processing covers the following sequence of steps, reported also in Figure 1:

1. *data import* from the heterogeneous databases (visits booking and municipal and parish archives) into a unified format;
2. *data cleaning and merge*, for data standardization and duplicates elimination with data merge;
3. *pedigree reconstruction*.

Step 1 is a sort of data structure standardization process, so that after the data import into a Unified Data Structure, each individual record has the same format, as shown in Figure 1. The unified structure is filled in by extracting and preprocessing the data coming from the historical archives, in order to reduce its imprecision and complete the partial information. This phase entails the definition of some rules for deriving new information from the available data. As an example, it is possible to reduce or completely define an individual birth date range, starting from the son's birth date, when the exact one is not available or it is completely omitted.

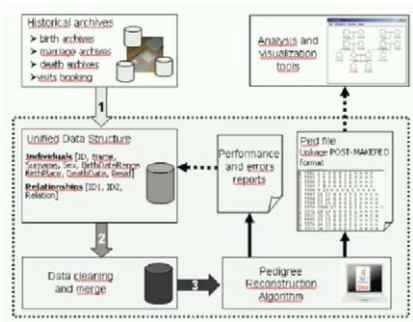


Figure 1- The overall strategy for data processing. The grey box highlights the procedures described in this paper

For this purpose, we distinguish the individuals reported in the archives into three categories: *Registered Persons, Relatives, and Spouses. Registered Persons* are the individuals identified trough their birth certificate or reported in the electronic municipal demographic archives (used also for visits booking). We assume that their personal identifying data are known without ambiguities (complete names, surnames and the precise birth date are always reported). *Relatives* are the persons reported in the different types of certificates (birth, marriage or death) and in the visits booking archive as parents or grandfathers of registered persons or spouses. Usually, such persons are identified only with their name and surname. Sometimes their age is reported, but frequently it is imprecise, having an error of ±3 years, while the birth date is omitted. *Spouses* are the individuals registered in the marriage certificates, whose demographic data (especially name and surname) are frequently different from those reported in birth certificates. This very frequent situation hampers the possibility of a strict match of a spouse with the corresponding registered person and forces us to resort to some additional strategies for record linkage performed in steps 2 and 3. The unified data structure allows the data cleaning and the reconstruction algorithm steps to be independent from the source databases structure, so that, if any modification of the source occurs or new information become available, after a suitable adjustment of the import procedure, it is necessary only to run again the algorithm without any other modification. On the other hand, the limit of this solution is that the reconstruction is not incremental with respect to new information inserted into the database.

Step 2 is fundamental to simplify the mechanisms of record linkage applied at Step 3. Its final goal is to eliminate possible duplicated individuals, by preserving and merging into a single record all the information available. If any information is reported in the duplicated records with different precision, it is necessary to resort to some a priori assump-

tion on the source archives reliability, in order to reduce the imprecision and the ambiguities during the merge. Before duplicates elimination, it is sometime necessary to perform names standardization for solving the problems in record linkage due to their different kind of variations.

Finally, in step 3 the reconstruction algorithm is performed in order to identify the parents of each registered person by exploiting all the available information derived at step 2. This phase requires to perform multiple record linkage, since the same person can be referred in several records (for example his birth certificate, all his sons birth certificates and his marriage certificate) with slightly different data. The strategy we adopt here is to perform different levels of checks between any couple of records, in order to detect all the relatives and, therefore, to correctly reconstruct the families composition. The final output is a *LINKAGE POST-MAKEPED* formatted file [11] containing the list of all the trios[1] of the population. This format has become a standard for most genetic analysis and visualization tools, so that the results need no more process to be evaluated and used. Moreover, some further output reports are provided about possible or proved data inconsistencies, which may be used to revise ambiguous, corrupted or missing information.

The following paragraphs describe in detail how steps 2 and 3 are performed. It is important to notice that such steps are completely independent from the original data sources and, after appropriate setup, can be applied also to different data sets.

Record linkage methods

In the different phases of the algorithm (data cleaning and pedigree reconstruction) we resort to record linkage procedures based on the match of the name, surname and birth date/age attributes. From our own experience, the most common problems in record linkage for pedigree reconstruction are generated by errors or imprecision in the attributes caused by *name variations*, *lack of precision in dates* (birth, marriage or death date), *incomplete information*, like partial name and/or surname, or availability of the individual age instead of his/her birth date. The more critical problem is represented by name variations, which cause failure to bring linkable records pairs together, due to different types of modifications on the name. The most frequent variations are: *spelling variations, phonetic variations, multiple and/or alternate names*.

- *Spelling variations*, usually due to data entry errors or ambiguity in demographic data, create a real problem which needs its own solution.

Example of name/surname variations are:
Bellazzi/Bellazi, Catterina/Caterina, Aluisa/Luisa.

- *Double or multiple names* which are very frequent in our data base. Typical situation is the one where the same person is referred in three archives with different names (which we call tokens):
  - birth certificate: *Anna Maria*
  - marriage certificate: *Anna*
  - son's birth certificate: *Maria*

If we should adopt an exact match of the names to link the records, a great number of relationships should be lost. To cope with this situation, we resort to a names comparison function based on the Levenshtein distance [12] applied to each token. Such function calculates a similarity score between two names taking into account: the number of tokens composing the name, the frequency of each token in the population (e.g. the name Maria is very frequent, so its contribution to the similarity score is lower than other less recurrent names) and the edit distance between each token. A threshold score is defined to accept or reject the match of two records based on the names.

- *Phonetics variations* result in significant modification of the name/surname usually due to the natural Italian language evolution and affect surnames coming from certifications registered before 1838. A dedicated strategy is needed for managing this kind of problem. At this time the idea is to identify (manually or automatically, using the edit distance) all the possible variations of every single name/surname, organize them into categories and associate each category to a "standardized" one (which typically is the one used since 1838). Once such categories have been defined, a preprocessing phase aimed to standardize all the names/surnames will be performed, before starting the actual reconstruction algorithm. Such strategy has not yet applied to real data, since currently only demographic data since 1838 are available in electronic form.
- Another frequent source of error is represented by imprecision in dates. This seems not so critical like the name's variation's problem. Our strategy here is to define some heuristic rules in order to derive a plausible range for the birth date of a person, by resorting to the fertile age concept. We assume that at the birth of a child the mother is never less than 14 or more than 60 years old, and the father is not less than 15 or more than 75 years old. This rule helps us to establish, with some other minor hints, a range for birth, death and marriage dates.

As the overall data processing cannot be completely automated, due to errors or imprecision caused by manual entries in the original archives, some reports of the ambiguous cases are provided to the archivists to manually check and correct the data in the source archives, as described in the following section.

An important step of the data cleaning process takes care of the elimination of duplicates and implements a strategy of data merge, which completes the information of the records with the data contained in the duplicates discarded.

The duplicates problem

Due to the source data organization, some individuals could have double or even three records (e.g. municipal and parish birth certificates and visits booking records), containing slightly different data (name, surname, birth date/age of an individual and of their parents). The issue at this step is to define a strategy for recognizing the duplicates and a series of rules to reduce the imprecision of the data during the merge. Such criteria establish, for example, which kind of record is considered the master and which are considered duplicates. This choice is driven by a priori

1 A *trio* is the numeric triplet that identifies an individual (called *proband*) together with his/her parents: *IDproband IDfather IDmother*

knowledge about the reliability and completeness of the different source archives. As an example, when we find duplicates in birth and municipal electronic archives, we maintain the second one, because its data are checked by the individual at the time of his/her visit and, therefore, is considered more reliable. In any case, we complete the information with the data coming from the birth certificate and, if the dates are reported with different precisions, we maintain the ones more accurate. A simple example of duplicates data is shown in Table 1, while Table 2 reports the resulting merged record.

Table 1 - The two original records

Archive type	Name	Surname	Date-Of-Birth-From	Date-Of-Birth-To
Municipal	Cristiana Francesca	Larizza	10-11-1956	10-11-1962
Birth	Cristiana Grazia	Larizza	10-11-1956	10-11-1957

Table 2 - The resulting record after duplicate elimination and data merge

Archive type	Name	Surname	Date-of-birth-from	Date-of-birth-to
Municipal	Cristiana Francesca Grazie	Larizza	10-11-1956	10-11-1957

The same kind of merge is performed also on data related to parents/relatives of each individual. In this case, it is very frequent having in one record the age and in another a birth date range (derived by exploiting the fertile age concept). In this case, for example, the mother/father age can be used to reduce his/her birth date range.

The reconstruction algorithm

The final goal of the pedigree reconstruction process is to create a trio for each registered person, by performing a series of cross checks between the information reported within the individual record (demographic data of her/his parents) and the marriage certificates of the potential parents. Such procedure can take to one of the following results: a) identification among the registered persons of the actual individual father/mother; b) creation of a fictitious father/mother that will be reused like any other registered person as possible parent of the remaining individuals.

In any case at the end of the procedure the relationship father-child/mother-child will be inserted into the corresponding table. More precisely, the algorithm covers the following steps, as shown in Figure 2.

1. Extraction from the data base of the list of registered persons potentially matching with the individual's father/mother. This phase is based on a blocking strategy which reduces sensitively the computational time in the comparison phase. The selected blocking attributes are: parent's sex, surname and birth date/age. Depending on the precision of the available data and on the extraction results, the blocking phase can be reiterated by relaxing some constraints in the birth date in order to try to obtain a not empty records list.
2. For every possible parent included in the list, a cross check, based on his/her marriage certificate, is performed in order to confirm the true match. It consists in verifying whether the potential father/mother is married with the mother/father declared in the son's birth certificate and in checking the compatibility of their ages.

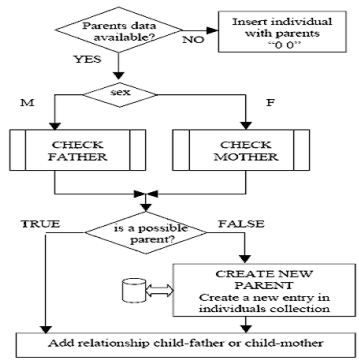


Figure 2 - Flow chart of the reconstruction algorithm.

3. If the marriage data confirm such parental relationship, the relation parent-child is created and stored in the corresponding table and the search stops.

If it is not found any possible parent among the registered persons of the list, including also fictitious individuals, (since the sequence of cross checks stops, due to failure of one of the matching conditions), or if the generated list is empty, the algorithm creates a further fictitious individual, based upon the data declared in the son's birth certificate and creates and stores the relation parent-child.

Results

After one a half year of project we collected and analyzed the data coming from municipal birth and marriage certificates registered since 1838 and municipal electronic archives of four towns available since 1985.

Table 3 - The data analyzed

Archive type	records #
Births	17.351
Municipal demography	3.480
Marriages	12.915
Individuals category	
Registered Persons (excluding duplicates)	19.139
Spouses	25.829
Relatives	75.529
Fictitious individuals	12.111

Table 3 reports the archive type and the number of records filled in into the Unified Data Structure during the import procedure from the historical archives. The total number of individuals whose pedigree has been derived is 31.164. The remaining records are used for cross checking between birth and marriage certificates. The data cleaning step detected 1692 records referring to duplicated individuals. Such records have been discarded by the reconstruction process, after addition or merge of the contained information into the remaining records.

The results of the pedigree reconstruction provided encouraging results being able to reconstruct a very big family of 10.634 individuals, corresponding to the 33.6% of the population analyzed. The total number of fictitious individuals created to complete the broken families is 12.111, but only 2.454 have been included into the big family. The remaining families are very small: we have 4.219 families whose dimension ranges from 3 to 117 individuals. Moreover, for some individuals it is possible to reconstruct the family up to seven generations. Also twins are in general correctly identified. The quality of the reconstruction algorithm was measured in terms of number of true positives n_m (number of relations correctly detected), number of false positives n_{fp} (number of multiple relations n_{mul} occurring when individuals are related to more than one father and/or mother) and number of false negatives n_{fn} (considered equal to the number of fictitious individuals $n_{fict,}$). Given N_m the number of expected relations (which is twice the number of individuals of the population) we can assume $n_m = N_m - n_{mul} - n_{fict}$. Given the previous assumptions, the *Recall* and *Precision* of the algorithm [10] were respectively estimated to be about 67.7% and 97.1%. At a first evaluation of the algorithm performance, it seems that a problem, although not very frequent, is the separation of a unique family into many family groups, whereas the aggregation of individuals belonging to different families into a single one never occurs. The debugging reports detected some conflicts in about 290 families' composition (spouses have a different number of sons), so our next goal is to identify and correct manually the errors causing this problem (at a first investigation they seem mainly due to transcription errors in the paper archives). On the basis of the debug performed until now, the number of records that will require a clerical manual review is about 3400. After the manual correction of the source archives, we expect to reduce the number of fictitious individuals and to unify some families into a single one. However, the results obtained until now are considered satisfactory by both the archivists and by the biologists that are able, on the basis of the pedigree obtained, to schedule the future visits and the genotyping of the most interesting individuals.

Conclusion

The automatic reconstruction of a population pedigree is a difficult problem that requires specific record linkage strategies. Various groups in the world are working on such techniques for various purposes (administrative, clinical, epidemiological studies, etc.), many of them being based on probabilistic approaches, but each methodology needs a customization based on the application context and the country (in particular, when names and surnames are chosen as linkage attributes) which heavily affects the performance (in terms of rate of errors) of the algorithm. In this paper we propose an approach that takes into account several context dependent ties on the linkage attributes in order to assess a true or non-true match between each record pair. The pedigree obtained automatically with the current algorithm seems satisfactory. However, to improve the results it will be necessary the manual intervention for detecting the causes of the errors in the tree and the correction of the data in the source archives. In the future we will make a more systematic evaluation of the matching rules, in order to tune the configuration parameters (for example, the threshold for the edit distance) and obtain an improvement in the algorithm performance. Moreover, we will evaluate the adequacy of the solution defined for managing the phonetic variations in names and surnames on real data.

References

[1] Martin-Sanchez F, Iakovidis I, Nørager S, Maojo V, De Groen P, Van der Lei J et al. Synergy between medical informatics and bioinformatics: facilitating genomic medicine for future health care. JBI, 2004 Feb; 37(1):30-42.

[2] Lander ES, Schork NJ. Genetic dissection of complex traits. Science. 1994 Sep 30; 265(5181):2037-48.

[3] Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nature Genetics, 2003 Mar;33:228-237.

[4] Cardon LR, Bell JI. Association study designs for complex diseases. Nature Genetics, 2001 Feb;2:91-99.

[5] Chapman NH, Thompson EA. Linkage disequilibrium mapping: the role of population history, size, and structure. Adv Genet. 2001;42:413-37.

[6] Mancosu G, Cosso M, Marras F, Borlino CC, Ledda G, Manias T et al. Browsing isolated population data. BMC Bioinformatics. 2005 Dec 1;6 Suppl 4:S17.

[7] Izzi C, Sanna-Cherchi S, Prati E, Belleri R, Remedio A, Tardanico R e al. Familial aggregation of primary glomerulonephritis in an Italian population isolate: Valtrompia study. Kidney Int. 2006 Mar;69(6):1033-40.

[8] Hamet P, Merlo E, Seda O, Broeckel U, Tremblay J, Kaldunski M et al. Quantitative founder-effect analysis of French Canadian families identifies specific loci contributing to metabolic phenotypes of hypertension. Am J Hum Genet. 2005 May;76(5):815-32. Epub 2005.

[9] Lin JP, Hirsch R, Jacobsson LT, Scott WW, Ma LD, Pillemer SR et al. Genealogy construction in a historically isolated population: application to genetic studies of rheumatoid arthritis in the Pima Indian. Genet Med 1999; 1(5): 187-93.

[10] Gu L, Baxter R, Vickers D, Rainsford C Record linkage: Current practice and future directions. Technical Report 03/83, April 2003, CSIRO Mathematical and Information Sciences, GPO Box 664, Canberra 2601, Australia.

[11] G. M. Lathrop, J. M. Lalouel, C. Julier, and J. Ott Strategies for Multilocus Linkage Analysis in Humans PNAS 1984;81:3443-6.

[12] V.I. Levenshtein, Insertions and Reversals, Cybernetics and control Theory 1966; 10(8):707-10.

Address for correspondence

cristiana.larizza@unipv.it

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Enhancing the Quality of Phylogenetic Analysis using Fuzzy Hidden Markov Model Alignments

Chrysa Collyda[a], Sotiris Diplaris[b], Pericles Mitkas[b], Nicos Maglaveras[a], Costas Pappas[a]

[a]*Lab of Medical Informatics, Faculty of Medicine, Aristotle University of Thessaloniki, Greece*
[b] *Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Greece*

Abstract

Any effective phylogeny inference based on molecular data begins by performing efficient multiple sequence alignments. So far, the Hidden Markov Model (HMM) method for multiple sequence alignment has been proved competitive to the classical deterministic algorithms with respect to phylogenetic analysis; nevertheless, its stochastic nature does not help it cope with the existing dependence among the sequence elements. This paper deals with phylogenetic analysis of protein and gene data using multiple sequence alignments produced by fuzzy profile Hidden Markov Models. Fuzzy profile HMMs are a novel type of profile HMMs based on fuzzy sets and fuzzy integrals, which generalize the classical stochastic HMM by relaxing its independence assumptions. In this paper, alignments produced by the fuzzy HMM model are used in phylogenetic analysis of protein data, enhancing the quality of phylogenetic trees. The new methodology is implemented in HPV virus phylogenetic inference. The results of the analysis are compared against those obtained by the classical profile HMM model and depict the superiority of the fuzzy profile HMM in this field.

Keywords:

bioinformatics, phylogenetic analysis, multiple sequence alignment, fuzzy integrals, Hidden Markov Models.

Introduction

The use of molecular data for inferring phylogenetic trees has gained considerable interest among biomedical researchers. Organisms, such as viruses do not leave fossil records, thus the only way to study their past is through the phylogenetic relationships of existing viruses. Phylogenetic analysis of protein and gene data can be accomplished by analyzing the genomic and proteomic sequence of the species. Many automatic methods have been developed for inferring phylogenetic trees, such as maximum parsimony[1], maximum likelihood[2] and distance methods[3]. The common feature of all these methods is that rates or patterns of change in sequences cannot be analyzed unless the sequences can be aligned [4], thus a robust multiple sequence alignment (MSA) is required as an input.

Bioinformatics offer a series of methods addressing this problem, such as CLUSTAL-W [5], PSI-BLAST [6] and HMMER [7] that can overpower classic methods of pairwise sequence alignment [8].

The well-known and widely used statistical method of characterizing the spectral properties of the residues of a genomic or proteomic pattern is the HMM approach. Profile HMMs have proved to offer a robust solution for MSA. Their wide use in bioinformatics has led to the creation of large profile databases [9],[10] that can offer biological knowledge (alignments, phylogenetic distribution, domain organization) for solving various problems, such as protein classification [11], building of phylogenetic trees [12], or gene function prediction.

However, an issue with the use of HMMs in MSA is its simplifying assumption on stochastic independence. This property, though, is not at all obvious when examining genomic or proteomic sequences; an underlying dependence may exist between current and previous states. Fuzzy HMMs have been introduced in speech recognition [13], in order to relax this assumption and resolve similar model definition issues, while in a previous paper we have introduced them in biological sequence analysis [14], by mathematically formulating the fuzzy profile Hidden Markov Model. Relative work has also been done in [15], where alignments from a fuzzy profile model were used for the description of the protein domain of kinases.

A characteristic of profile-HMMs, which have been used for MSA so far, is that these are finite models for the probability distribution over an infinite number of possible sequences. Profile-HMMs have the great benefit on generalized profiles that they are formally built on the probability theory. The disadvantage is that this theory restricts the flexibility of the models because the sum of the probability distribution over all modeled sequences must equal to one. In consequence, the probability of one sequence cannot be increased without decreasing the probability of another sequence in the profile-HMM. Fuzzy profile HMMs lack this restriction, thus they can be effectively used in order to better represent the sequences common residues and ultimately construct better phylogenetic trees.

In this paper, a new methodology of phylogenetic tree inference that makes use of the fuzzy profile HMM for multiple sequence aligning is presented. The fuzzy profile HMM representation, as we introduced it in [14], is defined by using fuzzy integrals and fuzzy operators in HMMs instead of probability theory. The classical HMM probabilities are replaced by fuzzy possibilities. The Choquet integral [16] is used for the integration over the HMM states and a new fuzzy measure is used for its application. Multiple sequence alignments are then obtained from the model and are used for the phylogenetic analysis of viruses coming for the HPV family. Phylogenetic analysis is finally performed with different algorithms and is evaluated using a bootstrapping schema.

The rest of the paper is structured as follows. First, the description of the experiment data is presented, while the following section describes in brief the methodology of constructing, training and acquiring alignments from the fuzzy HMM. In consecutive, we present the phylogenetic analysis schema, as well as the experimental results. Finally, we discuss the potential impact of the proposed methodology and end up with our conclusions.

Materials and methods

Data description

In the conducted experiments we used the E6 protein of several types, subtypes and variants of the Human Papillomavirus. Specifically, we acquired 78 different variants of the E6 protein using PSI-BLAST, coming from all known HPV types, as well as 30 protein sequences with homology to the HPV E6 protein, coming from various organisms. The latter were incorporated in the dataset in order to identify the discrimination capability of the different methodologies.

MSA with fuzzy profile Hidden Markov Models

Profile Hidden Markov Models

HMMs are statistical models used for MSA that allow the comparison of one gene or protein with a group of others, therefore facilitating the production of distinct differences between itself and the others. HMMs are a generalization of the profile in terms of statistical weights, rather than scores. At each position, the profile HMM gives the probability of finding a particular residue, an insertion, or a deletion. A profile HMM is composed by a number of interconnected states, each of which is able to emit observable output symbols, i.e. residues or gaps. Each state contains symbol emission probabilities and state transition probabilities. The symbol emission probabilities b_{jk} represent the probability of emitting each possible symbol k from a state j, whereas state transition probabilities a_{ij} are the probabilities of moving from the current state i to a next one j. An observation sequence $O=O_1O_2\ldots O_T$ can be generated by starting at an initial state and continuously changing of states, by also emitting symbols, until a specific end-state is reached at time T. The only visible outputs in this procedure are the emitted symbols, while the actual transition between states remains "hidden". Figure 1 depicts the structure of a profile HMM used for MSA, as introduced in [17]. Multiple alignments are used as a training set to build the model. One match state is assigned for each alignment column, insert states serve to insert extra symbols relative to the match states, while delete states allow for skipping positions in the training set aligned sequences. There are totally N=3*m+3 states in a profile-HMM, where m is the number of its match states.

The utilization of profile HMMs for MSA can be divided in three problems [18]:

- Problem 1: Computation of an observation probability according to the model: $P(O|\lambda)$. This is the problem of evaluation (HMM scoring), and it is usually solved using the forward-backward procedure.
- Problem 2: Computation of the state-sequence which fits the best to an observed sequence. This is the alignment problem. The Viterbi algorithm is usually used solve this problem and recover the hidden part of the model.
- Problem 3: Computation of the model parameters a_{ij}, b_{jk} and π to maximize the probability of one observation. This is the training problem, and the EM algorithm is usually exploited to this end.

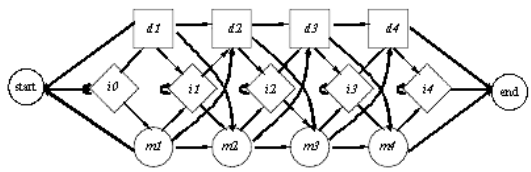


Figure 1- An example profile HMM with four match states.

MSA with fuzzy profile Hidden Markov Models

In a fuzzy profile Hidden Markov Model the classical HMM probabilities are replaced by fuzzy possibilities. Fuzzy integrals are used for aggregation over the states, while fuzzy operators are used instead of the algebraic ones. Though, the profile HMM structure is kept intact in terms of states and observations. The advantage of using fuzzy operators is that they are less constrained than classical integrals and probabilities, thus relaxing the independence assumptions that are necessary with probability functions in classical HMMs. This transformation also reduces the space of computations, thus yielding better response times.

In order to relax the additivity property of the classical HMMs constraint and take the relations between subsets into consideration, a generalization in terms of fuzzy measures has been introduced [19]. For the implementation of fuzzy HMMs, a possibility measure can be defined, such that

$$P(A \cup B) = \max(P(A), P(B)) \quad (1)$$

The max operator in the above equation is the fuzzy intersection operator. For a finite fuzzy set $X = \{x_1, x_2, \ldots x_n\}$, the density μ^i of a fuzzy measure μ can also be defined [20] as $\mu^i = \mu(\{x_i\})$, where $\mu(\{x_i\})$ represents a fuzzy measure, such

as the max operator, that is computed over the variables of the fuzzy set *X*.

This fuzzy measure can then be used with fuzzy integrals to compute integrations over fuzzy sets.

For the case of profile HMMs, the integration is done over the states that are a discrete set. In such cases, the discrete Choquet integral [21] can be used:

$$I = \sum_{i=1}^{n} \left[h(x_i) - h(x_{i-1}) \right] \mu_i^n \,, \tag{2}$$

with defined $x_i : 0 = h(x_0) \le h(x_1) \le \ldots \le h(x_n)$, and the discrete step μ_i^j:

$$\mu_i^j = \begin{cases} \mu\left(\{x_i, x_{i+1}, x_j\}\right) & \text{if } i \le j \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

The fuzzy profile HMM $\bar{\lambda} = (\bar{A}, \bar{B}, \bar{\pi})$ with *N* states $S=\{S_1,S_2,\ldots,S_N\}$ that can be observed through a space of observations Ω with observations $O=O_1O_2\ldots O_T$ corresponding to unknown state sequences $Q=q_1,q_2,\ldots, q_T$ can be fully defined by the matrices $\bar{A}$, $\bar{B}$ and $\bar{\pi}$, where $\bar{A}$ is the fuzzy state transition matrix, $\bar{B}$ is the fuzzy observation matrix and $\bar{\pi}$ is initial state fuzzy density. Two fuzzy variables $x \in X=\{x_1,x_2,\ldots, x_N\}$ and $y \in Y=\{y_1,y_2,\ldots, y_N\}$ are used to represent the state at time *t* and *t+1* [17].

In these terms, $\bar{\pi}_S(A)$ is the grade of certainty that the initial state is in A. Respectively, for $X_0 \subset X$ and $Y_0 \subset Y$, $\bar{\alpha}_Y(X_0 \mid Y_0)$ is the grade of certainty that the state at time *t+1* is in Y_0, given that the previous state was X_0. Concerning the observation space $\Omega_0 \subset \Omega$, $\bar{b}_i(\Omega_0)$ is the grade of certainty that the current observation is in Ω_0, given a current state S_i.

After defining the model we are able to address the three HMM problems. Specifically, the problem of HMM evaluation can be solved using the fuzzy forward-backward algorithm. $\bar{\alpha}_i(t)$ is the grade of certainty of $O=O_1O_2\ldots O_T$ and x_i at time *t*. The initialization step is $\bar{\alpha}_1(i) = \bar{\pi}_i \wedge \bar{b}_i(O_1)$, while the induction step becomes:

$$\bar{\alpha}_{t+1}(i) = \sum_{i=1}^{N} \bar{\alpha}_{ij} \left[\mu_i^n(t,j) - \mu_{i+1}^n(t,j) \right] \wedge \bar{b}_i(O_{t+1}) \tag{4}$$

where the sum is the discrete Choquet integral, the $\wedge$ operator stands for the fuzzy intersection operator, and μ_i^j is defined in Equation 3. From the above equation, it is possible to observe that the assumption of independence of the observation until time *t* is not necessary anymore neither is necessary the knowledge of the next state. The answer to the evaluation problem for the forward and backward variables respectively is:

$$P(O \mid \lambda) = \sum_{i=1}^{N} \bar{\alpha}_T(i) \,, \; P(O \mid \lambda) = \sum_{i=1}^{N} \bar{\beta}_1(i) * \bar{b}_i(O_1) \tag{5}$$

In the fuzzy case, the grade of certainty for a sequence is used to score the model. The Choquet integral is computed over the states at each time *t*, where the integration step $(\mu_i^n(t,j) - \mu_{i+1}^n(t,j))$ becomes a value *j* at time *t+1*.

Respectively, the fuzzy Viterbi algorithm, which is used for the alignment of new sequences to the model, uses the Choquet integral and multiplication for the fuzzy intersection operator in order to define the variable δ for the fuzzy case:

$$\bar{\delta}_t(i) = \max_{q_1,q_2,\ldots,q_{t-1}} \left\{ \bar{\pi}_{q_1} b_{q_1} \prod_{\tau=2}^{t} \left[\bar{\alpha}_{q_{\tau-1}q_\tau} \rho_\tau(q_{\tau-1}, q_\tau) \right] \bar{b}_{q_\tau}(O_\tau) \right\} \,, \tag{6}$$

where $\rho_t(i,j) = [\mu_i^n(t,j) - \mu_{i+1}^n(t,j)] / \bar{\alpha}_t(i)$. $\bar{\delta}_t(i)$ is the degree of certainty for a single state sequence finishing at time *t* in a state S_i.

Similarly, for training the fuzzy HMM model, the fuzzy version of the EM algorithm can be derived, again by using the fuzzy coefficient that multiplies the state transition coefficients and summing up using the Choquet integral.

Using fuzzy profile HMMs in phylogenetic analysis

After obtaining a multiple sequence alignment using the fuzzy Viterbi algorithm, the methodology of performing and evaluating phylogenetic analysis is depicted in Figure 2.

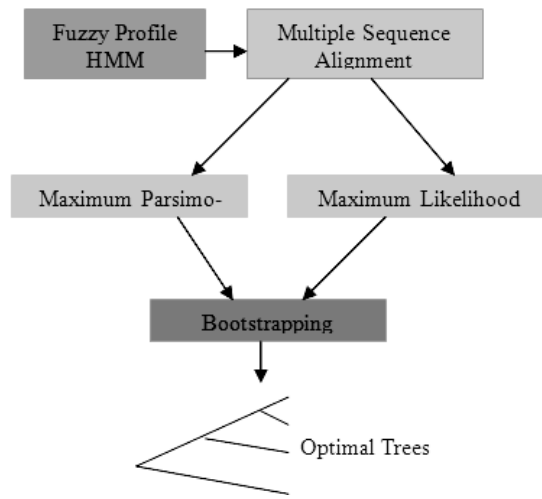


Figure 2 - Phylogenetic analysis using fuzzy HMM alignments

First, alignments are used to compute phylogenetic trees with the maximum parsimony and maximum likelihood methods.

The maximum parsimony method constructs trees on the basis of the minimum number of amino acid changes required to account for the data. This often results in the generation of hundreds of equally most parsimonious trees, making it difficult to justify the choice of a particular tree.

Similarly, for the maximum likelihood method, testing all possible trees is impossible, and it is also not computationally feasible to estimate the model for each tree. Therefore, the accepted strategy is to infer a "reasonable" tree topology with faster – although less reliable – reconstruction methods and use that tree to estimate the parameters.

By implementing a bootstrap analysis [22] of the produced trees though, it is possible to acquire a measure for the reliability of the alignments. Bootstrapping can approximate

the underline distribution by resampling from the original dataset and constructing a majority-rule consensus tree. Bootstrapping values can then be used as a confidence measure for the quality assessment of the alignments.

In the conducted experiments the abovementioned phylogenetic analysis schema was used for both classical profile HMM and fuzzy profile HMM alignments, and the confidence levels of the bootstrapping procedure were measured.

Results

In order to evaluate the performance of the fuzzy profile HMM alignments we have performed several comparative tests against the classical approach. The dataset was randomly divided in a 90%-10% manner into training and test set for the HMM models. Specifically, 97 sequences were used for training the model, while 11 sequences were used to create the multiple sequence alignment and consequently construct their phylogenetic tree. The test sequences came equally from the E6 protein HPV family and the homologous proteins obtained from PSI-BLAST. The classical profile HMM was then trained in 20 cycles, while the fuzzy profile HMM required 12 cycles for training. Alignments were obtained using the fuzzy Viterbi algorithm. Part of the resulting alignments yielding from the fuzzy profile HMM is depicted in Figure 3.

```
M-------L-M--L-A----M-M----E---LH-H----G----------|
D-VlIRC-Y-L-CH-KPL-C-EV----E--KVK-HIL-T-KAR-------
V-L----P-K---M-K---C-P------------HRL-E-----------
A-IsIDQ-KTT-SR-NPR-ST-VgtvtE-----iH--D-YLR--------
D-LsIRC-M-C-CL-KPL-S-PA----E--KLR-H-LN-SKRR-------
--------M-E-KL-QPK-T-I--------KN----C-S-QD--------
ElIfVGE-I-S-SS-KPA-I-SN------------IL--M-RCspgsssk
D-VlIRC-Y-L-CH-KPL-C-HV----E--KVR-HILD--KAR-------
Q-VrIRC-C-K-CH-KPL-S-PV----E--KTN-HIV--KKTQ-------
Q-LfMRC-Y-I-CH-KPL-S-WE----E--KEA-LLV-GNKR--------
--------YGD-YQ-SEYyG-P------gg-------N-YDFTpqgppss
```

Figure 3 - Part of alignment using fuzzy Profile HMM

In consequence, phylogenetic trees were inferred using the maximum likelihood (ML) and the maximum parsimony (MP) methods. Bootstrapping was also applied for each case. Table 1 illustrates log likelihood scores for the two cases of ML applications, while also contains the MP scores for each case. Optimal trees for the bootstrapped cases can be observed in Figures 4 and 5.

The fuzzy profile HMM was implemented in Java, partially using the BioJava API, while phylogenetic analysis was performed with the PHYLIP package [23].

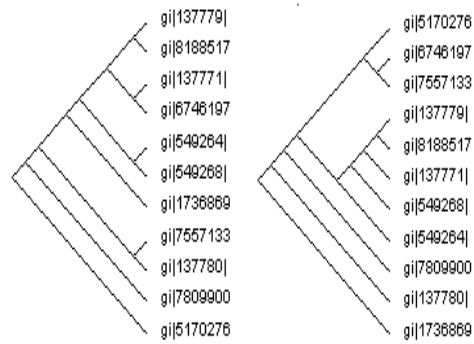


Figure 4 - Best trees for fuzzy and simple HMM by maximum likelihood method as calculated by the Phylip Package

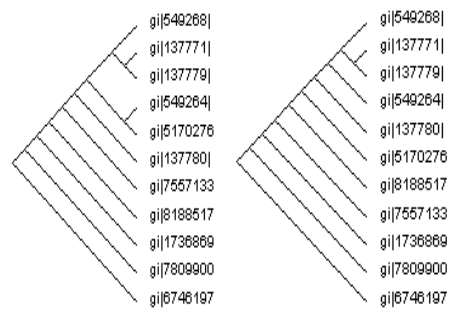


Figure 5 - Best trees for fuzzy and simple HMM by Maximum Parsimony method as calculated by the Phylip Package

Table 1 - Phylogenetic evaluation scores for different models and algorithms

Methods	HMM	FuzzyHMM
ML Simple	-13298.89	-13317.902
MP Simple	8507	8067
ML with Bootstrap	-13671.514	-13633.146
MP with Bootstrap	8609	8347

Discussion

By inspecting the inferred trees it is possible to observe that both models can discriminate between the E6 proteins and the rest of the homologous sequences. Concerning the evaluation scores, in the maximum-likelihood case, the tree topology, as well as the evaluation scores is similar for the two alignments in both simple and consensus trees. The maximum parsimony method for the fuzzy case yields models with lower values for both simple and bootstrapping methods, compared to the simple HMM alignment. This means that the fuzzy HMM can produce more parsimonious phylogenetic trees, a property that seems to come naturally from the non-independence assumption of the fuzzy profile HMM. Another issue worth pointing is the ability of the fuzzy HMM trainer to converge in less

cycles, thus building the model faster. In this sense, the computational cost for performing the training stage in the fuzzy case is reduced without sacrificing the quality of the alignments. Instead, the relaxation of the statistical independence assumption provides enhanced biological meaning to the alignments.

Conclusions

In this paper we presented a new methodology of phylogenetic tree inference that makes use of the fuzzy profile HMM for multiple sequence aligning. The fuzzy approach relaxes the independence restriction implied in classical profile HMMs, thus providing more biologically meaningful alignments. In terms of phylogenetic analysis this implies the constructions of more parsimonious trees in comparison with the classic HMM approach. Finally, we have shown this property by experimenting with HPV virus protein data. Future work involves the application of fuzzy HMM alignments in the creation of a whole new series of profiles that can then be used in protein classification.

Acknowledgments

This work was supported by the EPEAEK II project: "IRAKLEITOS: Scholarships targeted in basic research. Phylogenetic analysis of biological data".

References

[1] Fitch WM. Toward defining the course of evolution: Minimum change for a specific tree topology. Systematic Zoology 1971; 20:406-416.

[2] Felsenstein J. Evolutionary trees from DNA sequences: A maximum-likelihood approach. Journal of Molecular Evolution 1981; 17:368-376.

[3] Page RDM, Holmes EC. Molecular Evolution: A Phylogenetic Approach. Oxford: Blackwell Science, 1998.

[4] Salemi M, Vandamme AM, editors. The Phylogenetic Handbook. Cambridge University Press; 2003.

[5] Thompson D, Higgins DG, Gibson TJ. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22:4673-80.

[6] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Oxford d University Press Nucleic Acids Research 1997; 25(17):3389–3402.

[7] Eddy SR. Profile Hidden Markov Models. Bioinformatics 1998; 14:755-763.

[8] Pearson WR. Empirical statistical estimates for sequence similarity searches. In J. Mol. Biol. 1998; 276:71-84.

[9] Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer ELL, Studholme DJ, Yeats C, Eddy SR. The Pfam protein families database. Oxford University Press, Nucleic Acids research 2002; 30(1):276-280.

[10] Mulder NJ. InterPro. In: Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics, Jorde LB, Little PFR, Dunn MJ, Subramanian S, editors, John Wiley & Sons, 2435-2439; 2005.

[11] Diplaris S, Tsoumakas G, Mitkas PA, Vlahavas I. Protein classification with multiple algorithms. In Proceedings of the Tenth Panhellenic Conference on Informatics 2005, Volos, Greece. Bozanis P, Houstis EN (Eds), Springer, Lecture Notes in Computer Science; 3746:448-456.

[12] Collyda C, Maglaveras N, Pappas C. A Multiple Sequence Alignment Tool Using Hidden Markov Models For Phylogenetic Analysis. In Proceedings of the 4th European Symposium on Biomedical Engineering 2004; Patra, Greece.

[13] Tran D, Wagner M. Generalised Fuzzy Hidden Markov Models for Speech Recognition. In Proceedings of the International Conference on Fuzzy Systems 2002; Calcutta, India.

[14] Collyda C, Diplaris S, Mitkas PA, Maglaveras N, Pappas C. Fuzzy Hidden Markov Models: A New Approach In Multiple Sequence Alignment. 20th International Congress of the European Federation for Medical Informatics (MIE 2006), Maastricht, Netherlands. A. Hasman et al. (Eds.), IOS Press, Studies and Health Technology in Informatics; 124:99-104.

[15] Bidargaddi NP, Chetty M, Kamruzzaman J. Fuzzy Profile Hidden Markov Models for Protein Sequence Analysis. Proceedings of the IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology 2005; San Diego, CA, USA.

[16] Narukawa Y, Murofushi T, Sugeno T. Conditions for choquet integral representations. Proceedings of the Eighth International Fuzzy Systems Association World Congress 1999; 911–925.

[17] Krogh A, Brown M, Mian IS, Sjolander K, Haussler D. Hidden Markov models in computational biology: Applications to protein modeling. Journal of Molecular Biology 1994; 235:1501-1531.

[18] Rabiner A. A tutorial on Hidden Markov Models and selected applications in speech recognition. Proceedings of the IEEE 1989; 77(2):257-285.

[19] Wang Z, Klir G. Fuzzy measure theory. Plenum Press; 1992.

[20] Grabisch M, Murofushi T, Sugeno M, editors. Fuzzy measures and integrals : theory and applications. New York: Springer; 2000.

[21] Mohamed M, Gader P. Generalized hidden markov models. i. theoretical frameworks. IEEE Transactions on Fuzzy Systems 2000; 8(1):67–81.

[22] Felsenstein J. Numerical methods for inferring evolutionary trees. Systematic Zoology 1982; 27:783-791.

[23] Felsenstein J. PHYLIP: Phylogenetic Inference Package, Version 3.65. Seattle: Department of Genetics, University of Washington, 1993.

Address for correspondence

Lab of Medical Informatics, Faculty of Medicine, Aristotle University, GR54124, Thessaloniki, Greece

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Determining Transcription Factor Activity from Microarray Data using Bayesian Markov Chain Monte Carlo Sampling

Andrew V. Kossenkov[a], Aidan J. Peterson[b], Michael F. Ochs[c]

[a] *The Wistar Institute, Philadelphia, PA, USA*
[b] *Howard Hughes Medical Institute at the University of Minnesota, Minneapolis, MN, USA*
[c] *The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, MD, USA*

Abstract

Many biological processes rely on remodeling of the transcriptional response of cells through activation of transcription factors. Although determination of the activity level of transcription factors from microarray data can provide insight into developmental and disease processes, it requires careful analysis because of the multiple regulation of genes. We present a novel approach that handles both the assignment of genes to multiple patterns, as required by multiple regulation, and the linking of genes in prior probability distributions according to their known transcriptional regulators. We demonstrate the power of this approach in simulations and by application to yeast cell cycle and deletion mutant data. The results of simulations in the presence of increasing noise showed improved recovery of patterns in terms of χ^2 fit. Analysis of the yeast data led to improved inference of biologically meaningful groups in comparison to other techniques, as demonstrated with ROC analysis. The new algorithm provides an approach for estimating the levels of transcription factor activity from microarray data, and therefore provides insights into biological response.

Keywords:

Microarray analysis, Bayesian analysis, transcription factors

Introduction

The regulation of gene expression is a primary form of response in all cellular systems. This response is typically mediated by activation of transcription factors or complexes (both referred to here as TFs) that can induce or repress transcription of sets of genes by binding to upstream elements known as promoters. Identification of the activity level of individual TFs provides insight into biological processes activated or deactivated in specific samples. For instance, identification of upregulation of the ELK–1 TF can indicate activation of the oncogenic RAS-RAF pathway in a tumor cell, which is difficult to measure directly.

Microarrays provide insight into the global transcriptional response of cells, which should be useful for identification of TF activity. Although early studies using microarrays focused on "guilt-by-association" identification of genes that may function similarly to known genes [1] or biomarkers of disease [2], efforts have been made to use microarrays to link to transcription factor promoter sites [3].

These methods overlook a key aspect of transcriptional regulation, as they rely on clustering of genes into groups with each gene belonging to a single group. However, most, if not all, genes are likely to be multiply regulated, as evolution has been very effective in borrowing function by using existing genes in new roles. Even among genes regulated in the yeast cell cycle, only roughly 10% are associated with a single cell cycle phase [4]. This results in the identification of a large number of clusters with genes improperly grouped away from those involved in the same biological processes (e.g., in analyses of the yeast cell cycle data, a typical cluster analysis shows more than 20 clusters). This results in a significant loss of power for identification of TF activity.

Although our knowledge of transcriptional regulation is growing rapidly, in a recent study using Agilent human arrays, only ~1000 genes out of 20,000 were annotated with high reliability as to their TFs using TRANSFAC professional [5]. In order to recover the signal of TF activity, it is therefore highly desirable to maximize the signal by correctly grouping genes into multiple overlapping groups.

The problem of multiple regulation was identified reasonably early in the development of microarray technology. The application of singular value decomposition (SVD) to microarray data [6] addressed multiple regulation, however the orthogonality constraints led to less than ideal results, since biological processes are not independent. We took a different approach, applying our Bayesian Decomposition algorithm to microarray analysis [7]. This algorithm, described below, uses a series of constraints and a structure minimization argument to identify overlapping sets of genes.

A significant advantage to our approach is the ability to encode biological knowledge through prior probability distributions. In this work, we demonstrate how knowledge of coregulation through TFs can be encoded into the algorithm, leading to improved statistical power for the determination of the activity levels of biological processes.

Methods

In order to recover signatures of transcription factor activity, the analysis of microarray data will need to isolate patterns related to a biological process governed by a transcription factor, identify the genes associated with this pattern in the background of multiple regulation, and link these genes to transcriptional regulators. We will describe the Bayesian Decomposition (BD) algorithm in general, then show specifically the modifications that allow direct inference of transcriptional regulation to improve statistical power.

Bayesian decomposition

The fundamental factorization needed to identify overlapping groups of coexpressed genes is the recovery of a distribution matrix (**A**) and a pattern matrix (**P**) that multiply together to form a mock ("fitted") data matrix (**M**), which reproduces the data matrix (**D**) within the noise (). This relationship can be written as

$$\mathbf{D} = \mathbf{M} + \varepsilon = \mathbf{AP} + \varepsilon \,. \qquad (1)$$

For microarray data, **D** is generated from replicated experiments and therefore represents the best estimate for the expression of each gene in each condition. The computed matrices then provide the assignment of genes to patterns, **A**, and the assignment of conditions to patterns, **P**, as shown in Figure 1 for a hypothetical analysis of the cell cycle. In this example, the data is approximated by the multiplication of **A** and **P**, so that a gene (*N*) with complex behavior (transcribed strongly in G1 and weakly in G2), can have that behavior explained as a mixture of simpler behaviors (G1 and G2).

The factorization of **D** into **A** and **P** is generic, and as noted above, approaches using orthogonality criteria have been used. However, biological patterns will not be orthogonal, as this would imply independence. In fact, SVD applied to cell cycle data does not even readily recover phase signatures [6], while BD recovers signatures for the cell cycle phases as well as a signature for the entrained metabolic oscillator [7].

Probability distributions

BD implements a Markov chain Monte Carlo (MCMC) approach in order to solve Equation 1. The Markov chain uses a Gibbs sampler requiring relative probability measures between points in the distribution of possible solutions. These are provided according to Bayes' Equation,

$$p(\mathbf{A}, \mathbf{P}|\mathbf{D}) \mu\, p(\mathbf{D}|\mathbf{A}, \mathbf{P})\, p(\mathbf{A}, \mathbf{P}). \qquad (2)$$

The posterior probability, $p(\mathbf{A},\mathbf{P}|\mathbf{D})$, describes the probability of a model (**A** and **P**) given the data, and it is the distribution sampled by MCMC. The prior, $p(\mathbf{A},\mathbf{P})$, provides the probability of the model independent of the data. A simple example is that a model with negative copies of mRNA can be ruled out *a priori* and has zero prior probability. The likelihood, $p(\mathbf{D}|\mathbf{A},\mathbf{P})$, gives the probability that the data comes from the model, and this probability is related to the χ^2 distribution.

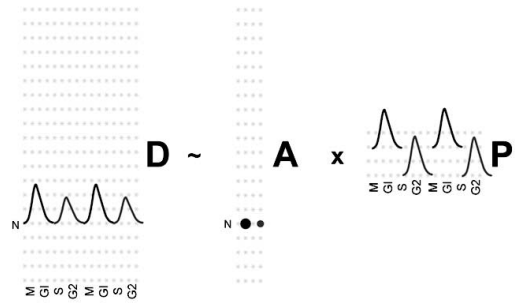


Figure 1 - The decomposition performed by BD

The prior encoded by the algorithm has three important features: positivity, correlated structure, and minimization. Positivity is incorporated by the inclusion of a one-dimensional atomic domain for each matrix in the model (**A** and **P**). In this domain, point masses (i.e., atoms) are created according to a prior distribution that is exponential in amplitude and uniform in location. The creation or destruction of atoms does not modify the prior distribution, which allows an approach to minimization of structure, as discussed below.

Correlations are introduced between points in the model (**A** and **P** matrices) by the mapping of atoms from the atomic domain to the matrices [8]. The mappings use kernel functions to spread the amplitude of each atom to one or more matrix elements, allowing linking of samples through correlations in **P** and linking of genes through correlations in **A**. We have used correlated structure in **P** to perform supervised learning [9], and here use correlations in **A** to provide prior information on TF regulation.

We use "birth and death" MCMC techniques for creating and destroying atoms. Since the prior distribution is unaffected by these actions, atoms can be eliminated readily, as long as the fit to the data is not adversely affected. This, coupled to internal mechanisms of amplitude exchange, leads to a minimization of structure (i.e., number of atoms). In the simplest application, this matches sparse matrix approaches.

Markov chain Monte Carlo sampling

The Markov chain begins with empty atomic domains, and thus empty **A** and **P** matrices. The algorithm attempts to birth atoms (created *ex vacuo*), move or exchange amplitude between atoms, and remove atoms in separate MCMC steps. Atoms are created according to the prior distribution and mapped through the kernel functions to the **A** and **P** matrices. The log likelihood is

$$\log L = \sum_i \sum_j \left\{ \frac{1}{2\sigma_{ij}} (D_{ij} - \sum A_{ip} P_{pj}) \right\}, \qquad (3)$$

so that changes in the likelihood can be easily calculated for any change in the matrices. The algorithm calculates this change in such a way as to allow resampling of amplitude to increase the speed of exploration of the posterior distribution.

After an equilibration period determined by the user, sampling of the distribution is done by recording atoms and mapping them to the **A** and **P** matrices. Statistical measures, such as the mean and standard deviation for each matrix element, can be calculated. The rows of the inferred **P** matrix (patterns) link conditions, while the columns of the inferred **A** matrix assign genes to each pattern. Convergence is checked by insuring that the χ^2 fit to the data is stable, and that multiple chains starting at random points reach the same solution.

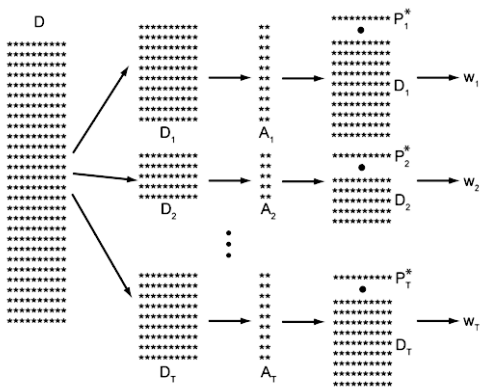


Figure 2 – Calculating weights for TF analysis

Estimating transcription factor activity

Previously we estimated TF activity without using prior information successfully for *S. cerevisiae*. We applied the original BD and subsequently linked genes associated with patterns to TFs [10]. However, for mammalian systems, a more statistically powerful approach will be required. The approach we propose here is to encode the knowledge of transcriptional regulation *a priori* during Markov chain sampling, thus borrowing power across genes.

Normalizing expression for each transcription factor

In order to map an atom to multiple matrix elements, it is necessary to determine the amount of the amplitude that should be devoted to each gene. This mapping cannot be made uniform across the matrix elements, as the overall copy number of mRNA produced will vary by gene and by biological process.

We address this issue by implementing a preprocessing step to determine weight vectors for each gene linked through a TF. The overall approach is shown in Figure 2. The data, **D**, are divided into T overlapping subsets, with each subset containing all genes regulated *a priori* by a given TF. Each subset, $\mathbf{D_t}$, is analyzed using the original BD without *a priori* correlations and positing $P+1$ patterns, where P is the total number of groups in T that contain any gene in $\mathbf{D_t}$. This provides for a pattern for each TF and a pattern for routine metabolic function, which BD typically isolates separately. The rows of the $\mathbf{A_t}$ matrices are normalized to unit amplitude, and the column from each with the lowest variance is taken to represent the pattern related to that TF. The dot product of the corresponding pattern, $\mathbf{P_t}^*$, with each row of $\mathbf{D_t}$ provides the weights, $\mathbf{w_t}$, for spreading an atom linked to this TF into the **A** matrix.

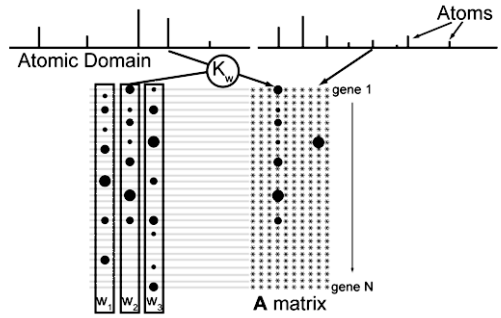


Figure 3 - Atomic domains and prior information

Sampling with prior information

We encoded TF information within BD by dividing the atomic domain related to the **A** matrix into two subdomains (see Figure 3). Atoms in the left half of the atomic domain spread to multiple elements in **A** through kernel functions, K_w, using the weight vectors, $\mathbf{w_t}$. The location of the atom determines the column of **A** and the weight vector $\mathbf{w_t}$ that are used. An atom in the right half maps to a single element, allowing prior information not supported by the data to be ignored.

As with BD, the new algorithm, BD-TF, starts with empty **A** and **P** matrices. As atoms are created, moved, or destroyed, changes are mapped to the matrices according to the scheme shown in Figure 3. If there is correlated structure in the data due to TF activity, a single atom will successfully recover this, and statistical power will be gained by use of K_w. Since the prior is defined on the atomic domain, the form is unchanged from the original BD algorithm. The calculated change in the log likelihood is affected by the correlation functions, K_w, but the likelihood function in Equation 3 includes summation across all matrix elements, so no change of form is needed. The equilibration and sampling proceed as in the original BD.

Analysis

We have analyzed three separate sets of data and compared our BD-TF results to BD and to some standard techniques. We first used simulations of the yeast cell cycle, which allowed us to increase noise levels to determine the behavior of the algorithm across many levels of noise. Second, we analyzed the widely studied yeast cell cycle data set [1], permitting comparison with other methods. Third, we analyzed the Rosetta compendium of yeast deletion mutants, which included error estimates for all data points [11]. An open problem in the field is the correct dimensionality for analysis (i.e., number of patterns). We have estimated this previously by multiple methods for the data used here, while, for the simulated data, we use the simulated dimensionality.

For the simulated yeast cell cycle data, we created expression levels for 288 genes at 48 time points over two cycles. The **P** matrix comprised 5 overlapping patterns - four that reflect cell cycle phases over two periods and one representing a metabolic oscillator with an amplitude 5% as large as the cell cycle pattern amplitudes and with twice the frequency [7]. The **A** matrix assigned genes to patterns in the expression profile, with most genes assigned to 2 – 4 patterns, reflecting the fact that in yeast only ~10% of cell cycle genes belong to a single phase [4]. Noise was added to the data matrix, including different levels of additive and multiplicative noise, using the widely accepted noise model [12]

$$D = N\left(0, \sigma_a\right) + e^{N(0,\sigma_m)} A_S P_S , \qquad (4)$$

where $\mathbf{A}_S$ and $\mathbf{P}_S$ are simulated amplitude and pattern matrices, and σ_a and σ_m are additive and multiplicative levels of noise respectively. Simulated data matrices with 154 different noise levels were created, varying σ_a from 0 to 6.5 and σ_m from 0 to 3. The data matrix without noise had maximum amplitude 3.15 and mean amplitude 0.65. We simulated 4 replicate arrays and calculated mean and standard deviations for each simulated element in each **D**.

The Cho data was analyzed using BD, as described previously [7], and using BD-TF. Groups of coregulated yeast genes were assembled based on literature reports of regulatory relationships between TFs and target genes, and only coregulation groups with at least five members were retained. In addition, we sought to enrich for true TF-target relationships by including only genes supported by evidence both of TF binding and of alteration in transcription when the TF was absent or overexpressed. For the cell cycle analysis, the regulators Mot3, Ndt80, Ste12, Swi5, Cbf1, Fkh1, Fkh2, Swi6, Mcm1, Swi4, and Rlm1 provided prior information with 5 – 16 target genes each.

In our previous work [7], we established that the best factorization used 6 dimensions (5 cell cycle phases, due to early and late G1 signatures as seen previously [4], and a metabolic oscillator signature). An ROC analysis was performed for BD by increasing the stringency of assignment of a gene to a pattern. Essentially, each gene had a mean value of its strength within a pattern and an uncertainty on that assignment from **A** based on MCMC sampling. By increasing the number of standard deviations away from zero required to assign a gene to a group, multiple estimates of the assignment of the genes to the patterns were made, allowing the ROC curve to be constructed. The gold standard for the analysis was based on the known molecular biology of gene coregulation independent of microarray studies, and this comprised 9 groups with 43 genes total [13]. The results were compared between hierarchical clustering, the original BD, and BD-TF.

The Rosetta compendium data was previously analyzed using BD at multiple dimensions, and consistency analysis determined that 15 dimensions were optimal [10]. The data was reanalyzed at 15 dimensions using BD-TF as for the yeast cell cycle data, except we included targets whose regulation by a TF was supported by a single type of experimental evidence. The set included the regulators Zap1, Ndt80, Mcm1, Gcn4, Dal80, Rtg1, Pdr1, Met4, Ume6, Ste12, Mot3, Gln3, Cbf1, Mig1, Rlm1, Msn4, and Msn2 with 5 – 19 target genes each. ROC analysis looked for recovery of the correlated groups. While this is somewhat circular, it indicated whether the prior information was being used appropriately by the algorithm. Unfortunately, the genes with known coregulation from cell cycle studies [13] were not in this data set, as they do not vary across the deletion mutants.

Results

We present our results in three sections: 1) simulations of cell cycle data, 2) analysis of the Cho data with ROC analysis, and 3) analysis of the Rosetta data with ROC analysis.

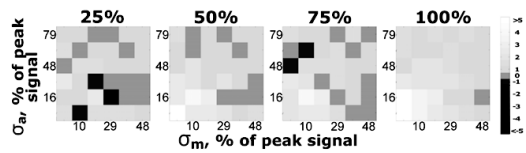


Figure 4 – χ^2 fits for simulation data

Simulations

In Figure 4, we show the $\log_2$ ratios of the χ^2 fits to the $\mathbf{A}_S$ matrix between BD-TF and BD. The heatmaps show the differences in the fits across different levels of noise and inclusion of prior information. In the figures, multiplicative noise increases to the right, from 0 to 48% of the peak signal; additive noise to the top, from 0 to 79% of peak signal; and the amount of prior information included (number of genes included from coregulation lists) increases as indicated by the percentages shown above each heatmap. Only the 36 simulations with the lowest noise levels are shown, as these cover levels of noise exceeding those in typical array experiments. Gray squares represent a neutral result (no improvement), while lighter squares represent improved fits with BD-TF and darker squares poorer fits. The advantage of using coregulation information increases as levels of prior information increase, as would be expected if the coregulation information is improving statistical inference by gaining power across genes.

Yeast cell cycle

In Figure 5 on the left, we present results of the application of BD-TF to the cell cycle data using ROC analysis based on known coregulation [13]. We compared the results using original BD (circles), BD-TF (squares), and shrinkage-based hierarchical clustering (triangles; performed previously [13]). BD-TF obtained an area under the curve of 0.82, compared with 0.83 for BD and 0.56 for hierarchical clustering. The lack of improvement from use of coregulation information reflected the lack of such data here, as we had TF data on only 67 of 788 genes, which was not adequate to improve inference over BD. However, we include the results to show the value in BD and BD-TF

that arises solely from the proper assignment of genes to multiple groups.

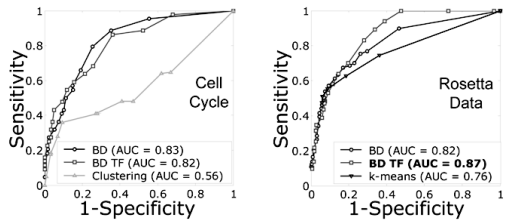


Figure 5 – ROC Curves for Cell Cycle and Rosetta Data

Rosetta compendium

The analysis of data from the Rosetta compendium data set using BD-TF (squares) was compared to K-means clustering (triangles) and the original BD analysis (circles) of the same data (Figure 5 right). Here, both the coregulation information and the "gold" standard gene lists were the same, so that the results demonstrated that the algorithm correctly used information about TF regulation and that such coregulation was reflected in the data. All techniques performed equally well at high specificity, however as sensitivity increased, BD-TF was superior due to a reduction of false positives with the inclusion of prior information on TF regulation.

Conclusion

This work demonstrates 1) the value of inclusion of prior knowledge on transcriptional regulation in the analysis of microarray data, and 2) the present limits of that knowledge. While the simulations showed a clear advantage in using this knowledge, especially at typical noise levels, the analysis of cell cycle data indicated that more prior information would be helpful. Nevertheless, the superiority of the BD-TF approach over clustering for microarray analysis is clear. Our knowledge of transcriptional regulation is rapidly increasing, and we expect improved statistical power with BD-TF over the next few years. This power will be critical to improved inference of biological process activity, especially with heterogeneous and limited samples typical in clinical settings. These samples introduce noise to an analysis focused on understanding biological response, such as in therapeutic interventions, and techniques to gain statistical power through use of existing biological knowledge will be critical to make progress.

Acknowledgements

This work was supported by NIH: LM008309 to mfo and CA06973 to mfo through the Sidney Kimmel Cancer Center.

References

[1] Cho RJ, Campbell MJ, Winzeler EA, Steinmetz L, Conway A, Wodicka L, et al. A genome-wide transcriptional analysis of the mitotic cell cycle. Mol Cell. 1998;2(1):65-73.

[2] Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000;403(6769):503-11.

[3] Vadigepalli R, Chakravarthula P, Zak DE, Schwaber JS, Gonye GE. PAINT: a promoter analysis and interaction network generation tool for gene regulatory network identification. Omics. 2003 Fall;7(3):235-52.

[4] Lukashin AV, Fuchs R. Analysis of temporal gene expression profiles: clustering by simulated annealing and determining the optimal number of clusters. Bioinformatics. 2001;17(5):405-14.

[5] Wingender E. TRANSFAC, TRANSPATH and CYTOMER as starting points for an ontology of regulatory networks. In silico biology. 2004;4(1):55-61.

[6] Alter O, Brown PO, Botstein D. Singular value decomposition for genome-wide expression data processing and modeling. Proc Natl Acad Sci U S A. 2000;97(18):10101-6.

[7] Moloshok TD, Klevecz RR, Grant JD, Manion FJ, Speier WFt, Ochs MF. Application of Bayesian Decomposition for analysing microarray data. Bioinformatics. 2002 Apr;18(4):566-75.

[8] Sibisi S, Skilling J. Prior distributions on measure space. Journal of the Royal Statistical Society, B. 1997;59(1): 217 - 35.

[9] Ochs MF, Moloshok TD, Bidaut G, Toby G. Bayesian Decomposition: Analyzing microarray data within a biological context. Annals of the New York Academy of Sciences. 2004;1020:212 - 26.

[10] Bidaut G, Suhre K, Claverie JM, Ochs MF. Determination of strongly overlapping signaling activity from microarray data. BMC Bioinformatics. 2006 Feb 28;7(1):99.

[11] Hughes TR, Marton MJ, Jones AR, Roberts CJ, Stoughton R, Armour CD, et al. Functional discovery via a compendium of expression profiles. Cell. 2000;102(1): 109-26.

[12] Rocke DM, Durbin B. A model for measurement error for gene expression arrays. J Comput Biol. 2001;8:557 - 69.

[13] Cherepinsky V, Feng J, Rejali M, Mishra B. Shrinkage-based similarity metric for cluster analysis of microarray data. Proc Natl Acad Sci U S A. 2003 Aug 19;100(17): 9668-73.

Address for correspondence

Michael F. Ochs,
Oncology Biostatistics and Bioinformatics,
The Sidney Kimmel Comprehensive Cancer Center
at Johns Hopkins, 550 N.
Broadway, Suite 1103,
Baltimore, MD 21205, USA
Email: mfo@jhu.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Ensemble Stump Classifiers and Gene Expression Signatures in Lung Cancer

Lewis Frey[a1], Mary Edgerton[b], Douglas Fisher[c] and Shawn Levy[d]

[a] *Biomedical Informatics Department, University of Utah, USA*
[b] *Department of Anatomic Pathology, M.D. Anderson Cancer Center, USA*
[c] *Department of Electrical Engineering and Computer Science, Vanderbilt University, USA*
[d] *Department Biomedical Informatics, Vanderbilt University, USA*

Abstract

Microarray data sets for cancer tumor tissue generally have very few samples, each sample having thousands of probes (i.e., continuous variables). The sparsity of samples makes it difficult for machine learning techniques to discover probes relevant to the classification of tumor tissue. By combining data from different platforms (i.e., data sources), data sparsity is reduced, but this typically requires normalizing data from the different platforms, which can be non-trivial. This paper proposes a variant on the idea of ensemble learners to circumvent the need for normalization. To facilitate comprehension we build ensembles of very simple classifiers known as decision stumps – decision trees of one test each. The Ensemble Stump Classifier (ESC) identifies an mRNA signature having three probes and high accuracy for distinguishing between adenocarcinoma and squamous cell carcinoma of the lung across four data sets. In terms of accuracy, ESC outperforms a decision tree classifier on all four data sets, outperforms ensemble decision trees on three data sets, and simple stump classifiers on two data sets.

Keywords:

microarray, decision trees, ensembles, stumps

Introduction

Methods for finding robust mRNA signatures of cancer that remain consistent across experiments and microarray platforms (i.e., oligonucleotide and cDNA) have remained elusive in the bioinformatics literature. From a machine learning perspective this is expected, since many microarray data sets have a scarcity of sample (e.g., a few hundred); moreover, each sample has thousands of probes (i.e., continuous variables) resulting in a very pronounced curse of dimensionality. With thousands of variables from which to choose, the constructed classifier can overfit the specific data and cannot generalize to other data sets. It poses a challenge when applying machine learning techniques to discover a set of relevant probes that constitute a robust mRNA signature for the cancer. There are a number of papers describing the pitfalls of overfitting expression data and the failure of some classification models to do better than chance [1,2,3].

By combining data from different platforms, problems of data sparsity and overfitting can be mitigated. The microarray data sets available in repositories are growing at a rapid rate. The creation of data sharing initiatives such as oncomine.org and the Cancer Biomedical Informatics Grid (caBIG™) enable the combination of multiple data sets to find better classifiers.

However, combining data across platforms is challenging. First of all, there are multiple microarray platforms and these can differ in the types of probes arrayed (i.e. variables measured) for the specimens. A method for mapping the probes across platforms is first required to relate the results. The different means of getting expression levels from the platforms need to be consistently compared, which often requires normalizing the expression levels between platforms.

A methodology that supports the generation of classifiers that find easily interpretable, robust mRNA signatures of cancer that *generalize across experiments and platforms* is needed. By robust mRNA signatures what is meant is a set of probes that are consistently associates with a cancer type. This paper describes a novel method for combining data sets to discover classifiers that use robust mRNA signatures which generalize across experiments and platforms. This is done by using a classifier to focus on a limited number of predictive probes in the data that persist across data sets. *The result of our approach is a distinct classifier for each data set under investigation, but importantly the construction of each such classifier is informed by all the available data.*

Empirical results by Holte [4] showed that simple, single test classifier trees, referred to as *Stumps*, can be surprisingly close in accuracy to more complex decision tree classifiers[5] in many of the domains tested. Results on data sets used in machine learning have shown that combining multiple classifiers boosts classification accuracy by creating variance among the constituent classifiers. Our approach, Ensemble Stump Classifier (ESC), is a kind of subspace sampling. It combines different probes that have slightly dissimilar classification of the sample using a majority vote to boost accuracy due to the variance in the sample they accurately classify. Using Stump classifiers as the "base" classifier of the ensemble supports the creation of simple ensemble classifiers.

1 L.F. was supported in part by a fellowship from the National Library of Medicine (5T15LM007450). M.F. and D.F. were supported in part by a National Institute of Health grant (1R01LM008000) to M.E.

Materials and methods

Four classifier-methods -- decision trees (C5.0), ensemble decision trees, Stump rule and ESC -- are applied to the data sets. Ten-fold cross validation is used to derive the accuracy measure for the data sets. Each data set is divided into 10 mutually-exclusive folds, consisting of 10% of the data. One of the 10-folds is held out as a testing set and the remaining 90% is used as the training set. Then a different fold is selected as the testing set and the remaining 90% serves as the training set. The average testing accuracy is calculated across all 10-folds.

C5.0, ensemble C5.0 and stumps

C5.0 is a commercial implementation of Quinlan's decision tree approach [5]. The ensemble decision tree classifier is computed using C5.0's implementation of the boosting approach of Freund and Schapire [6]. Stumps are single variable-test decision trees. Each Stump is made up of one probe at the root and has the best threshold for the training set that separates the two classes. Stumps are obtained by running C5.0 rules on a single data set with pruning set to a 1% confidence level with only single variable rules allowed.

Ensemble stump classifier (ESC)

The ESC is a new classifier approach that bypasses the need to normalize across multiple data sets. The base classifier for the ensemble is the Stump classifier. *The outcome of our method is a separate classifier for each data set, but notably each classifier is instructed by all the available data.*

Basic representations

ESC's form of ensemble learning is different than boosting. The idea that learning is occurring over multiple data sets is key. To illustrate, consider three artificially constructed objects, one from each of three data sets (see Table 1a).

Probes P_{1j} (i.e., a *matching probe set*) represent "synonymous" probes across the three data sets. The three probes -- P_{11}, P_{12}, and P_{13} -- can be thought of as the same probe with different scales and labels. The probe P_{11} has an expression value for the probe in match Probe Set 1 and Data Set 1.

The ESC algorithm repeatedly finds the "probe" (though with different labels and scales) that is "collectively" best over all data sets. This step leads to one decision stump per data set, each of which uses the same decision variable/probe. For example, if probe P_1 were found to be "collectively best" over three data sets, then the three decision stumps for each data set would have the form expressed in Table 1b.

The threshold for match Probe Set 1 in Data Set 1 is equal to 0.7 in Table 1b. Given an expression value for probe P_{11} equal to or below 0.7 the Stump rule predicts Class 1, otherwise Class 2. For Probe Set 1 and Data Set 2 the threshold is different than Data Set 1 (compare threshold 0.7 for P_{11} and threshold 10 for P_{12}). Even though the probes are the same, the thresholds can be different between data sets. This occurs because the data sets are not normalized to each other. The threshold is only consistent within a data set.

Building an ESC

The following are the steps to generate a full ensemble stump classifier (as in Table 1b):

1. Determine which probes match each other and construct rows based on the matching (i.e., "synonymous") probes across the data sets. This was briefly described in the basic representations section, and is described in more detail in the match probes section.
2. Partition each of the three data sets into folds so that there are training and testing sets. In our experimental design we use 10-fold cross validation.
3. Learn the stump rules, one for each training set and matching probe set, from training folds. This was illustrated in the basic representations section (e.g., each P_{ij} in Table 1b), and follows the same process given for stump rules.
4. Reorder the stumps across the data sets based on the probes that are "collectively" best overall.
5. Determine the best number of probes to be in the ESC based on training accuracy.

Step 4 is implemented by ordering the stump rules generated in Step 3 according to the weighted average training accuracy across the data sets. The average accuracy is weighted by the size of the data sets. Step 5 is implemented by incrementally adding stump classifiers in order of quality (accuracy) until one fails to improve the accuracy of the ESC.

Classification with an ESC

Table 1a illustrates the data structures for three data sets with probes having expression values and each instance having a true class assigned. The true classes are assigned to the instance through an expert classifying the instance. For cancer tissues the expert is often a pathologist. Table 1b illustrates the data structure for an ESC consisting of three probes over three data sets. The thresholds and predictive classes are shown for the probes for each data set.

Table 1c is the substitution of the expression values for the probes from Table 1a into the data structures for ESC in Table 1b. The predicted classes are present for each probe. The cells that are grey in Table 1c are incorrect class assignments given the expression values and the thresholds. The bolded classes are correct assignments. For example in Table 1c, P_{11} is incorrectly assigned while P_{12} and P_{13} are correctly assigned. Two out of three correct gives an accuracy of 66.7%.

The ESC majority vote classifier is obtained by going down a column and choosing the majority class predicted by the rules. An odd number of probes are used in the ensemble for simplicity and to ensure there is always a winner. For example in Data Set 1 the majority class is 1 because P_{21} and P_{31} are both correctly assigned to Class 1. For Data Set 2 the Majority Class is 2 because P_{12} and P_{22} are both correctly assigned to Class 2. Going across the majority vote row gives three out of three correct for an accuracy of 100%. Table 2 is the accuracy of the probe sets across the data sets and the ESC accuracy. Because we have not normalized the classifiers in ESC, the column or data set against which the test instance is compared must be known; we will address this limitation in the discussion.

Table 1 - Obtaining accuracy given an instance and an example ESC.

(a) Test instance with expression values and class.			
	Data 1	**Data 2**	**Data 3**
Test Instance with expression values for probes	(P_{11}=0.8, P_{21}= 0.85, P_{31}=1.5)	(P_{12}=12, P_{22}=14, P_{32}=15.5)	(P_{13}=0.02, P_{23}=0.35, P_{33}= 0.55)
True Class	Class 1	Class 2	Class 1
(b) Example ESCs for the three probe sets and data sets:			
Match Probe	**Data Set 1**	**Data Set 2**	**Data Set 3**
P_{1j}	P_{11} <= 0.7 ->Class 1 P_{11} > 0.7 -> Class 2	P_{12} <= 10 ->Class 1 P_{12} > 10 -> Class 2	P_{13} <= 0.1 ->Class 1 P_{13}> 0.1 -> Class 2
P_{2j}	P_{21} <= 0.9 ->Class 1 P_{21} > 0.9 -> Class 2	P_{22} <= 12 ->Class 1 P_{22} > 12 -> Class 2	P_{23} <= 0.3 ->Class 1 P_{23} > 0.3 -> Class 2
P_{3j}	P_{31} <= 1.2 ->Class 2 P_{31} > 1.2 -> Class 1	P_{32} <= 15 ->Class 2 P_{32} > 15 -> Class 1	P_{33} <= 0.5 ->Class 2 P_{33} > 0.5 -> Class 1
ESC	Majority class predicted by P_{11}, P_{21} & P_{31} Stump rules	Majority class predicted by P_{12}, P_{22} & P_{32} Stump rules	Majority class predicted by P_{13}, P_{23} & P_{33} Stump rules
(c) Class predictions given expression values from 1(a) and classifiers from 1(b). Bold is correct and grey is incorrect as compared to true class in 1(a).			
Match Probe	**Data Set 1**	**Data Set 2**	**Data Set 3**
P_{1j}	0.8 > 0.7 -> Class 2	12 > 10 -> **Class 2**	0.02 <= 0.1 ->**Class 1**
P_{2j}	0.85 <= 0.9 ->**Class 1**	14 > 12 -> **Class 2**	0.35 > 0.3 -> Class 2
P_{3j}	1.5 > 1.2 -> **Class 1**	15.5 > 15 -> Class 1	0.55 > 0.5 -> **Class 1**
ESC	Majority class predicted by P_{11}, P_{21} & P_{31} Stump rules -> **Class 1**	Majority class predicted by P_{12}, P_{22} & P_{32} Stump rules -> **Class 2**	Majority class predicted by P_{13}, P_{23} & P_{33} Stump rules -> **Class 1**

Table 2 - Accuracy across the three data sets of the match probe sets and the ESC.

Match Probe	P_{1j}	P_{2j}	P_{3j}	ESC
Accuracy	66.7%	66.7%	66.7%	100%

Gene expression data

There are four data sets. Collectively, these data sets are made up of two affymetrix arrays and two cDNA microarrays data sets. There are two types of tumor tissue: adenocarcinoma and squamous cell carcinoma of the lung. The Su et al. [7] data set consists of 28 samples with 14 adenocarcinomas and 14 squamous cell carcinomas of the lung with 12,533 affymetrix probes (i.e., continuous variables). The Bhattacharjee et al. [8] data set consists of 160 samples with 139 adenocarcinomas and 21 squamous cell carcinomas of the lung with 12,600 affymetrix probes. The Yamagata et al. [9] data set consists of 20 samples with 9 adenocarcinomas and 11 squamous cell carcinomas of the lung with 4417 cDNA probes. The Garber et al. [10] data set consists of 52 samples with 39 adenocarcinomas and 13 squamous cell carcinomas of the lung with 24,192 cDNA probes.

Matched probes

Before obtaining an ESC we must identify matching probes across data sets. While the other classifier methods that we have described will be applied to each of the four data sets independently, an ESC will be learned from the "combined" data set that we are about to describe.

A probe corresponds to part of a gene. Thus, different probes can be associated with the same gene. The same gene may be referenced by different probes in different data sets obtained on different platforms. In order to use an ensemble method that generalizes across different data sets, the Affymetrix and cDNAs probes must be mapped to each other. For this paper all probes are matched via Affymetrix U95A probe names. For the ensemble method the probe sets are matched as follows. Bhattacharjee et al. and Su et al. are joined via their almost identical chips, which are

U95Av2 and U95A Affymetrix chips respectively. Thus they have total overlap of 12533 unique probes. Garber et al. and Yamagata et al. also have similar cDNA probes with a unique overlapping set of 2106 accession id probes.

Consequently, we have two pairs of similar platforms. It is trivial to map probes across platforms within each pair. To map across the pairs (i.e., across all four platforms) an online resource called ProbeMatchDB [11] is used to map these 2106 accession ids onto the U95A Affymetrix ids in a many-to-many mapping. The many-to-many mapping occurs because some of the 2106 cDNA probes have multiple U95A Af-fymetrix ids associated with them. All possible combinations of probes from the four data sets with matching U95A Af-fymetrix ids are used to construct the probe sets. This match-ing process results in the creation of 4491 probe sets. The probe sets are made up of four probes with one from each data set. A probe set can be thought of as a set of references for the same gene. In the discussion that follows we use "probe set" and "gene" synonymously.

Results

The four classification methods described above (i.e., ESC, Stump Rule, C5.0 and Ensemble C5.0) are compared using 10-fold cross validation testing accuracy, standard error and the average number of variables (Avg Var) for each classifier (See Table 3). Each C5.0 tree, Ensemble C5.0 forest and Stump Rule has been built in a way that is informed by a single data set. Thus, we will speak of classifications made by these approaches as using probes.

The ESC method uses the same 4,491 match probe sets or genes for each data set. As we have noted, the construction of ESCs are informed by multiple data sets, but to classify a datum with an ESC we must know the data set (e.g., what column in Table 1) that test instance is drawn.

The ESC does better than C5.0 on all four data sets, better than ensemble C5.0 on three out of four sets and better on two out of four for Stump Rule. Note that the ESC uses probes that are robust across the four data sets rather than using the best probes for each given data set. The Stump Rule, C5.0 and ensemble C5.0 are all using the best probes within a given data set, which gives them an advantage for computing accuracy, although not on generalization to other data sets.

A measure of the complexity of the classifiers is given in the mean number of variables across folds for a classifier. The ESC converges to a classifier of three variables (genes: BPAG1, KRT5 and ABCC5) for nine out of the 10-folds. As explained earlier, this number of genes does not result from a user-defined threshold, but to add more stumps (i.e., genes) to the ensemble would reduce training set accuracy. In the one remaining fold ESC converges to BPAG1, KRT5 and SIAT7B which results in a total of four variables used by the classifier. The thresholds in the rules are stable within each data set and hence there is consensus among the stumps in the cross-validation step Ensemble C5.0, C5.0 and Stump use varying number of probes for each data set, but they do not find probes that generalize across all four data sets.

Discussion

ESCs are not limited to genomic data per se, and we are interested in their characteristics from a machine learning standpoint. Ensembles boost accuracy by insuring variability in classification behavior among the base classifiers. In bagging [12] this variance stems from bootstrap sampling and the instability of the classifiers that are constructed with these differing sample. In random subspace selection the requisite variance comes from selection of differing variables [13] on which to form the classifier. To some extent such designs might be motivated by a desire to use off-the-shelf, greedy decision tree induction, which is a standard base classifier of ensemble approaches. Another way that variance could be achieved would be to modify classifier systems to directly return a set of sufficiently good and sufficiently different classifiers. In fact, this is our approach, though our base classifiers are stumps – we incrementally add "best" stumps (as assessed across multiple data sets) until performance drops.

The limitation of knowing what column in ESC to apply to a particular data set is moot within a lab or facility since they will always compare on the data they collect. The benefit comes from being informed by data collected at other labs and facilities. In practice it can be used to identify signature genes that when combined enable the construction of high accuracy classifiers, even in hold-out data sets.

Table 3 - Adenocarcinoma vs. Squamous Cell Carcinoma for 10-fold cross validation test accuracy for ESC, Stump classifiers, C5.0 decision tree, Ensemble C5.0 and SVMs with standard error (SE) and average number of variables across folds (Avg Var) in the classifier.

Data Set	ESC			Stump Rule			C5.0			Ensemble C5.0		
	Avg Acc %	SE	Avg Var	Avg Acc %	SE	Avg Var	Avg Acc %	SE	Avg Var	Avg Acc %	SE	Avg Var
Bhattacharjee	95.6	1.3	3	96.9	1.0	1	93.7	1.6	3	94.4	1.7	1.3
Su	93.3	4.4	3	83.3	7.5	1	91.7	5.7	1	90.0	7.1	1
Garber	91.8	3.4	3	92.7	3.1	1	87.0	4.8	1.4	88.3	4.4	1.2
Yamagata	95.0	5.0	3	40.0	10.0	1	80.0	11.1	1	90.0	6.7	1

Whenever ensembles succeed in boosting accuracy it can be argued that the base classifiers, by definition, must be overfitting or underfitting the data. The ensemble then results in a classifier that moves towards a "best" fit to the data. In the case of an ensemble of decision trees that boost accuracy, the move is probably from overfit to better fit. Overfitting may account for cases where C5.0 underperformed Stump Rule on

some data sets. In the case of ESCs the move is probably from underfit to better fit. ESCs are not only very simple, are formed from stumps that generalize well across data sets, and as a result may underfit any given data set.

One reason stumps that generalize well across data sets are desirable with microarray data is that artifacts can be introduced into the data by the lab that collects the data. Thus, the very best stump for a particular data set may exploit an environmental peculiarity of the lab that is collecting the data. When dealing in thousands of probes, any small laboratory bias may systematically influence the value of one or more probes [14]. Problems with data or facility bias are found in other contexts. For example, Evans and Fisher [15] found that a feature (i.e., printing press) that was highly predictive of a printing defect in a particular printing plant provided no insights to why similar problems occurred at other plants.

ESC builds classifiers that generalize across data sets (labs, facilities), and informs data collectors about probes that generalize beyond their data collection processes. When ESCs learned across data sets are contrasted with the best classifiers within a data set, our methodology can also point out lab biases that should be remedied.

ESCs were motivated initially by the desire to combine evidence from multiple sources of gene expression data, thereby mitigating the curse of dimensionality. We are interested then in what ESCs can tell in a biological domain. Notably, the ensemble method identifies the genes KRT5, BPAG1 and ABCC5 as informative across the four data sets that we examined. Importantly, the ESC method gives convergent support for the relevance of these probe sets relative to the findings in the original studies. Using hierarchical clustering Bhattacharjee et al. [8] found KRT5 and BPAG1 to be highly expressed in squamous cell carcinoma. Using hierarchical clustering Yamagata et al. [9] confirmed in their data that KRT5 and BPAG1 are highly expressed in squamous cell carcinoma. Using hierarchical clustering Garber et al. [10] also identified KRT5 and BPAG1 as highly expressed in squamous cell carcinoma. Using a SVM ranking method Su e al. [12] did not identify high expression of KRT5 or BPAG1 as predictive of squamous cell carcinoma. However, they did identify ABCC5 as predictive of squamous cell carcinoma. The convergent findings of these alternative methods provide additional support of the utility of the ESC method.

Conclusion

The ESC method does well with only three variables. These results suggest the existence of compact sets of genes with single thresholds, which can be measured using multiple modalities that consistently and accurately predict diagnosis. The building of data repositories and data exchange standards such as oncomine.com and the caBIG™ can assist in the discovering other robust mRNA signatures of cancer using ESC. The ESC finds in a greedy fashion the best available matched probes that can be used in an mRNA signature that generalizes across data sets.

References

[1] Simon, R. Diagnostic and prognostic prediction using gene expression profiles in high-dimensional microarray data. British Journal of Cancer 2003: 89; pp. 1599-1604.

[2] Simon, R., Radmacher, M.D., Dobbin, K. and McShane, L.M. Pittfalls in the Use of DNA Microarray Data for Diagnostic and Prognostic Classification. J. Natural Cancer Institute 2003: 95(1); pp. 14-18.

[3] Michiels, S., Koscielny, S. and Hill, C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. Lancet 2005: 365; pp. 488-92.

[4] Holte, R. Very Simple Classification Rules Perform Well on Most Commonly Used Datasets. Machine Learning 1993: 11; pp. 63-91.

[5] Quinlan, J.R. Induction of decision trees. Machine Learning 1987: 1; pp. 81-106.

[6] Freund, Y. and Schapire, R.E. Experiments with a new boosting algorithm. In Machine Learning: Proceedings of the Thirteenth International Conference 1996; pp. 148 - 156.

[7] Su, A. et al. Molecular Classification of Human Carcinomas by Use of Gene Expression Signatures.* *Cancer Research* 2001: 61; pp. 7388-7393.

[8] Bhattacharjee, A, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. PNAS: 2001: 98 (24); pp. 13790-13795.

[9] Yamagata N. et al. A training-testing approach to the molecular classification of resected non-small cell lung cancer. Clin Cancer Res 2003: 9; pp. 4695-704.

[10] Garber, M.E., et al. Diversity of gene expression in adenocarcinoma of the lung. PNAS 2001: 98 (24); pp. 13790-13795.

[11] Wang,P. et al. (2002) ProbeMatchDB—a web database for finding equivalent probes across microarray platforms and species. Bioinformatics, 18, pp. 488–489.

[12] Breiman, L. Bagging Predictors. Machine Learning 1996: 24(2); pp 123-140.

[13] Breiman, L. Random Forests. Machine Learning 2001: 45(1); pp. 5-32.

[14] Shi, Leming et al. The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nature Biotechnology 2006: 24(9); pp. 1151 – 1161.

[15] Evans, B. and Fisher, D. Overcoming Process Delays with Decision Tree Induction 1994: 9(1); pp. 60-66.

Address for correspondence

Lewis Frey, Ph.D.
26 South 2000 East; Salt Lake City, UT 84112; USA

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

From "Glycosyltransferase" to "Congenital Muscular Dystrophy": Integrating Knowledge from NCBI Entrez Gene and the Gene Ontology

Satya S. Sahoo[1], Kelly Zeng[2], Olivier Bodenreider[2], Amit Sheth[1]

[1]Kno.e.sis Center, Department of Computer Science and Engineering, Wright State University, Dayton, OH, USA
[2]U.S. National Library of Medicine, NIH, Bethesda, Maryland, USA

Abstract

Entrez Gene (EG), Online Mendelian Inheritance in Man (OMIM) and the Gene Ontology (GO) are three complementary knowledge resources that can be used to correlate genomic data with disease information. However, bridging between genotype and phenotype through these resources currently requires manual effort or the development of customized software. In this paper, we argue that integrating EG and GO provides a robust and flexible solution to this problem. We demonstrate how the Resource Description Framework (RDF) developed for the Semantic Web can be used to represent and integrate these resources and enable seamless access to them as a unified resource. We illustrate the effectiveness of our approach by answering a real-world biomedical query linking a specific molecular function, glycosyltransferase, to the disorder congenital muscular dystrophy.

Keywords:

knowledge integration, Semantic Web, RDF, Entrez Gene, Gene Ontology

Introduction

A common scenario in biomedical research involves the correlation of genomic data with disease information, in other words, associating genotype and phenotype information. In the particular scenario illustrated in this paper, a researcher is interested in glycosylation and its implications for one disorder: congenital muscular dystrophy. The biological process of glycosylation results in the post-translational addition of glycosyl groups (saccharides) to proteins (and lipids). Various enzymes, namely glycosyltransferases, catalyze glycosylation reactions.

From the functional annotation of gene products with terms from the Gene Ontology (GO), a researcher can identify the genes having the molecular function of catalyzing the transfer of specific glycosyl groups (e.g., *hexosyltransferase*, for hexosyl groups). Known associations between these genes and diseases can then be mined from resources such as NCBI's Entrez Gene (EG), where phenotypic information is recorded as pointers to the Online Mendelian Inheritance in Man (OMIM) knowledge base [3]. (See the Materials section for a presentation of GO and EG.)

In order to validate the hypothesis of possible association between the molecular function *glycosyltransferase* and the disease *congenital muscular dystrophy*, a researcher could simply search EG for the term *glycosyltransferase*, and all records containing the string "glycosyltransferase" in GO annotations would be returned. This approach, however, is suboptimal for at least two reasons. First, the term *glycosyltransferase* might appear as a substring in other GO terms (e.g., in *UDP-glycosyltransferase*), possibly leading to false positives. Conversely, not all GO terms related to *glycosyltransferase* actually contain the string "glycosyltransferase" (e.g., *acetylglucosaminyltransferase*, a kind of *glycosyltransferase*), possibly leading to false negatives.

To avoid false positives and false negatives, a careful researcher would likely start exploring the Gene Ontology database to create a list of *glycosyltransferase*-related terms by selecting the term *glycosyltransferase* itself (GO:0016757) and all its descendants, including specialized types of *glycosyltransferase*, such as *acetylglucosaminyltransferase*. This researcher would then look for the genes annotated with any of the *glycosyltransferase*-related terms. Resources such as the web browser AmiGO [1] support such searches and can retrieve the genes associated with any descendant of a given GO term. Finally, each of the genes found associated with any of the *glycosyltransferase*-related terms must be searched individually in EG, looking for mentions of the disease *congenital muscular dystrophy* (as an OMIM phenotype) in the corresponding records.

The procedure described above is evidently inefficient, time consuming and error prone as several web interfaces need to be utilized (AmiGO and Entrez), and as the results of the search in one resource need to be copied and pasted as search terms in the other. The main reason for such inefficiency is that high quality resources such as GO and EG have been designed primarily for consultation by humans, not for automated processing by agents or integration in applications. Moreover, these resources have been developed by different groups, independently of each other and are therefore not interoperable. No system currently supports complex queries such as: *Find all the genes annotated with glycosyltransferase-related terms in GO and associated with the disease congenital muscular*

dystrophy in OMIM. Typically, querying across the different knowledge sources is accomplished manually through meticulous work or requires the development of complex and customized software applications.

In this paper, we propose an integrative approach to querying across knowledge sources. More specifically, we have applied Resource Description Framework (RDF) [4] standard developed by the World Wide Web Consortium (W3C) to integrate knowledge from GO and EG, and used this integrated resource to answer complex queries. We use the scenario presented earlier to illustrate the advantages of this approach. This work is a pilot contribution to the *Biomedical Knowledge Repository* under development at the U.S National Library of Medicine (NLM) as part of the *Advanced Library Services* project [8]. This repository integrates knowledge not only from structured resources (database and knowledge bases), but also from the biomedical literature (e.g., MEDLINE), in order to support applications, including knowledge discovery.

Background

Information integration is one of the most challenging areas of research in Computer Science [11]. The use of heterogeneous schemas for data storage, that are designed primarily to ensure optimization of storage space, makes it extremely difficult for users to query data sources in an integrated manner. (The interested reader is referred to [12] for a survey of approaches to information integration.) The Semantic Web provides a common framework that enables the integration, sharing and reuse of data from multiple sources. The use of a representation formalism based on a formal language enables software applications to 'understand' and reason over information. Recent research in Semantic Web technologies has delivered promising results to enable information integration across heterogeneous knowledge sources.

The Resource Description Framework (RDF) is a W3C-recommended framework for representing data in a common format that captures the logical structure of the data. This is in contrast to pure storage aspects addressed by traditional relational database schema. The RDF representational model uses a single schema in contrast to multiple heterogeneous schemas or Data Type Definitions (DTD) used to represent data in XML by different sources. Hence in conjunction with a single Uniform Resource Identifier (URI), all data represented in RDF form a single knowledge repository that may be queried as one knowledge resource. An RDF repository consists of a set of assertions or triples. Each triple is constituted of three entities namely, the *subject* – the triple pertains to this entity, the *object* – the entity that states something about the object and the *predicate* – the relationship between the *subject* and the *object*. For example, as shown in Figure 1, assertions such as *acetylglucosaminyltransferase* (GO:0008375) is a kind of *hexosyltransferase* (GO:0016758) and the gene *LARGE* (EG:9215) has molecular function *acetylglucosaminyltransferase* (GO:0008375) can be represented as RDF triples.

The RDF triples often share nodes, thus forming a graph. For example, the two triples shown in Figure 2 share the node *acetylglucosaminyltransferase* (GO:0008375). The resulting graph is shown in Figure 2. The graph structure created by RDF is key to information integration in the Semantic Web.

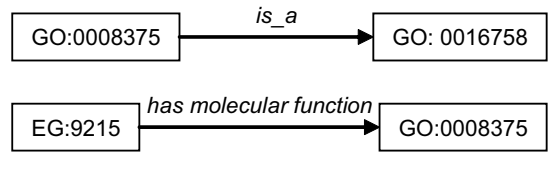
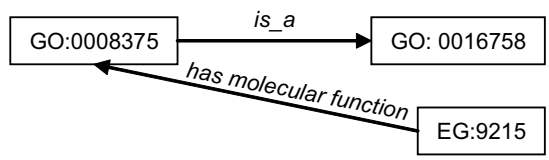


Figure 1 - Example of RDF triples



Fiigure 2 - Example of RDF graph

Materials

The **Gene Ontology** (GO) seeks to provide a consistent description of gene products [13]. GO consists of three controlled vocabularies for biological processes (9,234 terms), molecular functions (7,456 terms) and cellular components (1,804 terms). The GO monthly releases are made available on the GO website in various formats, including RDF. The version of GO used in this study is dated of September 2006.

The **Entrez Gene** (EG) database records gene-related information from sequenced genomes and of model organisms that are focus of active research [9], totaling about two million genes. EG contains gene information about genomic maps, sequences, homology, and protein expression among others [9]. In contrast to GO, EG is not available in RDF, but in XML (converted from ASN1 by the program *gene2xml* provided by NCBI), and can be downloaded from the NCBI website. The version of EG used in this study is dated of July 2006.

Methods

Our integration method can be summarized as follows and is illustrated in Figure 3. First, we extract manageable subsets from the two resources to be integrated. We then have to convert the EG subset from XML to RDF. Finally, we load both RDF resources in a common store, apply inference rules, and issue queries against it.

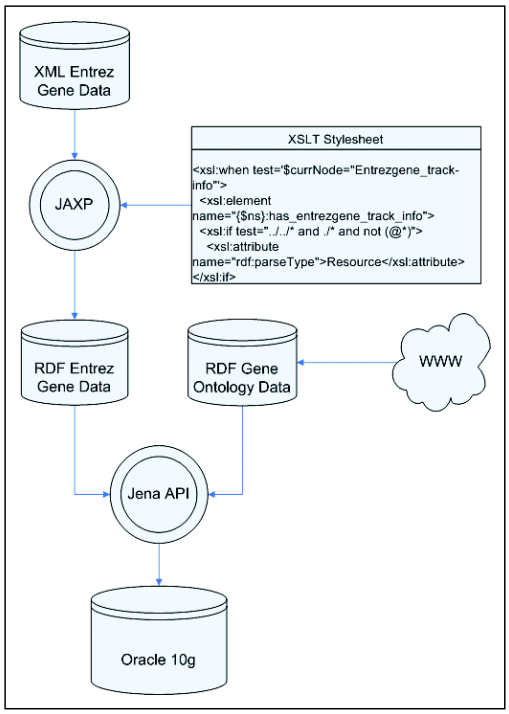


Figure 3 - Overview of the integration method

Creating subsets

The entire Entrez Gene data file (in XML format) is very large (50 GB) and unnecessarily difficult to manipulate. In order to obtain a manageable subset from EG, we restricted the gene records to two species: *Homo sapiens* (human) and *Mus musculus* (mouse). The resulting EG subset contains a total of 99,861 complete gene records (excluding obsolete records).

Converting XML format Entrez Gene data to RDF

A key element of our integration approach is the conversion of Entrez Gene from XML to RDF. There are many issues involved in the conversion of XML data into RDF format, including modeling the original semantics of the data, filtering redundant XML element tags, linking data entities using meaningful named relationships and identifying entities consistently within and across resources. Unlike traditional XML to XML conversion, XML to RDF conversion should exploit the advantages of the RDF model in representing the logical structure of the information.

We chose not to convert the element tags of the native EG XML representation mechanically into the *predicates* of the RDF triples. Instead, we manually converted the XML element tags into meaningful and standardized relationship names that convey explicitly the semantics of the connection between the *subject* and the *object*. For example, the element *<Org-ref_taxname>* was mapped to the more meaningful relationship named *has_source_organism_taxonomic_name*.

We selected the eXtensible Stylesheet Language Transformation (XSLT) [6] for converting the EG XML information into RDF, because this approach allows for a clean separation between the application (using Java API for XML Processing (JAXP)) and the conversion logic (using XSLT stylesheet). Once the stylesheet is created, it can serve as an auxiliary file for existing programs realizing the XML to RDF conversion. In other words, the major interest of this approach is that no specific code is required for the conversion, because the transformation logic resides entirely in the stylesheet.

Loading the two resources into a single data store

Some of the requirements for our RDF store include native support for the RDF graph data model, support for persistence and indexing of the RDF triples, support for extensive collections of triples, and availability of a query language for the RDF graph. After surveying available RDF storage solutions, we decided to use Oracle Spatial 10g [7] as the RDF storage system.

The RDF file resulting from the XSLT conversion of the original XML file for EG and the downloaded RDF version of GO are both loaded into a single RDF store. More precisely, the RDF resources are first converted to the NTriple format using the Jena API [10] and loaded into the RDF database using a utility program provided by Oracle.

Applying inference rules

Unlike the Web Ontology language OWL, RDF provides no direct support for inference. However, inference rules can be implemented in the RDF store to make explicit the semantics of some predicates. For example, the relationships *is_a* and *part_of* used in GO are partial order relations, thus being reflexive, antisymmetric and transitive. The inference rules we created for implementing the transitivity and combination of these two relationships are shown in Table 1. The inference rules are stored in a rule base created in Oracle 10g.

Table 1 - Inference rules for is_a and part_of in GO

Relation	is_a	part_of
is_a	IF <**x** *is_a* **y**> & <**y** *is_a* **z**> THEN <**x** *is_a* **z**>	IF <**x** *is_a* **y**> & <**y** *part_of* **z**> THEN <**x** *part_of* **z**>
part_of	IF <**x** *part_of* **y**> & <**y** *is_a* **z**> THEN <**x** *part_of* **z**>	IF <**x** *part_of* **y**> & <**y** *part_of* **z**> THEN <**x** *part_of* **z**>

Querying the RDF Graph with SPARQL

SPARQL [5] is a query language for RDF graphs, equivalent to SQL, the Structured Query Language, for relational databases. Unlike SQL, SPARQL does not require users to be familiar with the data model (e.g., tables, foreign keys), but simply to indicate how entities of interest relate to each other. For example, the structure of the query: *Find all the genes annotated with the GO molecular function glycosyl-*

*transferase (*GO:0016757) *or any of its descendants and associated with any form of congenital muscular dystrophy* is represented in Figure 4.

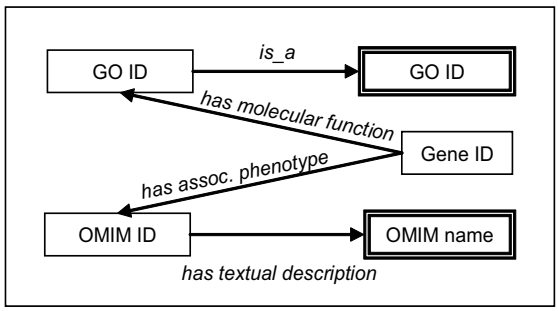


Figure 4 - RDF graph corresponding to the query above

```
SELECT distinct t,g,d
FROM TABLE(SDO_RDF_MATCH(
'(?t is_a GO:0016757)
 (?g has molecular function ?t)
 (?g has_associated_phenotype ?b2)
 (?b2 has_textual_description ?d)',
SDO_RDF_Models('entrez_gene'),
SDO_RDF_Rulebases('entrez_gene_rb'),
SDO_RDF_Aliases(SDO_RDF_Alias('','')), null) )
where (
    REGEXP_LIKE(LOWER(d), '((.)*(congenital)(.)*)')
AND REGEXP_LIKE(LOWER(d), '((.)*(muscular)(.)*)')
AND REGEXP_LIKE(LOWER(d), '((.)*(dystrophy)(.)*)'));
```

Figure 5 - Example of SPARQL query (simplified)

The query can be understood as finding a path in the RDF graph using a predetermined set of semantic relationships and would be formulated as follows. Because of the inference rules implementing the transitivity and reflexivity of the *is_a* relationship, the condition on the GO annotation "*glycosyltransferase (*GO:0016757) *or any of its descendants*" is easily expressed by '`?t is_a GO:0016757`'. The link between genes and GO terms is expressed by '`?g has_molecular_function ?t`'. Similarly, the link between genes and OMIM diseases is expressed by '`?g has_associated_phenotype ?b2`' (OMIM ID) and '`?b2 has_textual_description ?d`' (disease name). Finally, direct constraints are put on the GO term on the one hand ('`?t is_a GO:0016757`', to select *glycosyltransferase* (GO:0016757)) and on disease names on the other (where a regular expression is used to select disease names containing the strings "congenital", "muscular" and "dystrophy"). The actual (but simplified) SPARQL query is shown in Figure 5.

Results

One integrated RDF repository for Entrez Gene and GO

The subset of Entrez Gene restricted to *Homo sapiens* (human) and *Mus musculus* (mouse) as biological sources comprises 99,861 gene records. Once converted to RDF, it consists of 772,530 triples. The RDF version of GO contains 293,798 triples. Overall, there are over one million triples in the store created for this experiment, which is relatively small in comparison to the 411 million triples resulting from the conversion of the entire EG to RDF [2].

Biological query result: extended example

The SPARQL query presented above returned one result, corresponding to one path in the graph between the GO term *glycosyltransferase (*GO:0016757) and OMIM disease names containing (variants of) the string "*congenital muscular dystrophy*".

This path involved the human gene *LARGE like-glycosyltransferase* (EG:9215), annotated with the GO term *acetylglucosaminyltransferase* (GO:0008375), a descendant of *glycosyltransferase (*GO:0016757). Also involved in this path is the OMIM disease identified by MIM:608840. The name (textual description) of this disease is *Muscular dystrophy, congenital, type 1D* and contains the required substrings "congenital", "muscular" and "dystrophy". The instantiated RDF graph with path between *glycosyltransferase* (GO:0016757) and *Muscular dystrophy, congenital, type 1D* is shown in Figure 6.

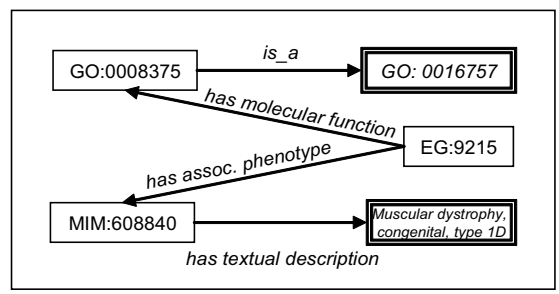


Figure 6 - Instantiated RDF graph

This simple SPARQL query provides an easy way of testing the biological hypothesis under investigation, i.e., the existence of a possible link between glycosylation and *congenital muscular dystrophy*. On manual inspection of the Entrez Gene record, we also note that the given gene may be involved in the development and progression of meningioma through modification of ganglioside composition and other glycosylated molecules in tumor cells.

Discussion

Significance

In this study, we demonstrated the feasibility of integrating two biomedical knowledge resources through RDF. We also provided anecdotal evidence for the benefits of such integration by showing how *glycosyltransferase* can be linked to *congenital muscular dystrophy*. The integrated resource is greater than the sum of its parts as it supports complex queries that could typically not be handled otherwise without tedious manual intervention or customized software applications.

Integrated resources based on a graph model are particularly important in an exploratory context where researchers need to "connect the dots" in order to validate an hypothesis. This approach also facilitates intuitive

hypothesis formulation and refinement. For example, after verifying that glycosyltransferase is linked to congenital muscular dystrophy, our researchers may narrow the focus of their wet lab experiments to only hexosyltransferase out of the potential seven glycosyltransferases. Analogously, they can focus their research on Muscular dystrophy, congenital, type 1D, out of several other diseases.

Arguably, the graph data model of RDF resources is more intuitive than the database schemas. In fact, the RDF data model enables us to model the inherent logical relations between entities that mirror the human cognitive model of the real world. Additionally, the RDF data model offers more flexibility than database schemas for accommodating changes to the underlying model.

Generalization

The integration approach demonstrated in this study can be generalized to more complex queries and to additional information sources. For example, many additional constraints can be easily added to the query presented earlier by exploiting other properties represented in GO or EG. Examples of such constraints include restricting the annotations to specific evidence codes (e.g., *TAS*) and narrowing the query to a specific model organism.

Only two resources are currently integrated in our RDF store. However, this approach can be generalized to other resources including pathway databases, microarray resources, disease ontologies and virtually all the structured knowledge bases currently under the umbrella of the Entrez system, including UniGene and HomoloGene. Knowledge extracted from unstructured sources such as the biomedical literature can also be integrated. Creating such an extensive repository of biomedical knowledge is one of the goals of the *Advanced Library Services* project under development at NLM.

Unresolved issues and challenges

In addition to scalability issues, which can be addressed by mature software and the next generation of hardware, challenges include the identification and organization of entities and relationships. Heterogeneous resources can interoperate in a RDF graph only if the entities shared by these resources are identified consistently. The namespace provided by the UMLS is expected to play an important role for the permanent identification of biomedical entities. In contrast to entities for which organizational schemes currently exist (terminologies and ontologies), the named relationships used to connect data entities during the conversion of EG from XML to RDF are currently not formalized in an ontology of relationships. As a consequence, only limited reasoning can be supported by the RDF graph. As sizeable ontologies of relationships become available, they too will be used for normalizing knowledge in our repository. RDF schemas and OWL will also be investigated.

Acknowledgments

This research was supported in part by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM) and by the Integrated Technology Resource for Biomedical Glycomics (5 P41 RR18502), funded by the National Institutes of Health National Center for Research Resources.

References

[1} AmiGO: Gene Ontology browser [cited 11/29/06; Available from: http://www.godatabase.org/]

[2] BioRDF subgroup: Health Care and Life Sciences interest group [cited 11/29/06; Available from: http://esw.w3.org/topic/ HCLSIG_BioRDF_Subgroup]

[3] Online Mendelian Inheritance in Man, OMIM (TM). McKusick-Nathans Institute for Genetic Medicine, Johns Hopkins University (Baltimore, MD) and National Center for Biotechnology Information, National Library of Medicine [cited 11/29/06; Available from: http://www.ncbi.nlm.nih.gov/omim/]

[4] Resource Description Framework (RDF), [cited 11/29/06; Available from: http://www.w3.org/TR/2004/REC-rdf-primer-20040210/]

[5] SPARQL Query Language for RDF. W3C Working Draft 2006 [cited 11/29/06; Available from: http://www.w3.org/TR/rdf-sparql-query.]

[6] XML Schema Language Transformation (XSLT) [cited 11/29/06; Available from: http://www.w3.org/TR/xslt]

[7] Alexander, N., Ravada S., "RDF Object Type and Reification in Oracle"—Technical White Paper [cited 11/29/06; Available from: http://download-east.oracle.com/otndocs/tech/semantic_web/pdf/rdf_reification.pdf]

[8] Bodenreider O, Rindflesch TC. Advanced library services: Developing a biomedical knowledge repository to support advanced information management applications. Technical report. Bethesda, Maryland: Lister Hill National Center for Biomedical Communications, National Library of Medicine; September 14, 2006.

[9] Maglott D, Ostell J, Pruitt KD, Tatusova T. "Entrez Gene: gene-centered information at NCBI", Nucleic Acids Res. 2005 January 1; 33(Database Issue): D54–D58.

[10] McBride, B.. Jena: A Semantic Web Toolkit. IEEE Internet Computing 2002;6, 6 (Nov. 2002), 55-59.

[11] Sheth AP. "Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics", in Interoperating Geographic Information Systems. M. F. Goodchild, M. J. Egenhofer, R. Fegeas, and C. A. Kottman (eds.), Kluwer, Academic Publishers, 1999, pp. 5-30.

[12] Shvaiko P, Euzenat J. 2005. A survey of schema-based matching approaches. Journal on Data Semantics 4: 146-71.

[13] The Gene Ontology Consortium. Gene Ontology: tool for the unification of biology. Nature Genetics 2000;25: 25-29.

Address for correspondence

Olivier Bodenreider, National Library of Medicine
8600 Rockville Pike, MS 3841, Bethesda, MD 20894, USA.
Email: olivier@nlm.nih.gov. Phone: (301) 435-3246.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Building a Research Model for Human Genetic Variation Knowledge Management

Yulong Gu[a], James Warren[a]

[a] *Department of Computer Science, University of Auckland, Auckland, New Zealand*

Abstract

Organizational knowledge management (KM) research studies the nature of knowledge, the scope of KM, the factors and mechanisms that affect KM outcomes, as well as theoretical KM frameworks. This paper discusses the implications of past studies for the KM efforts in the human genetic variation (HGV) research domain and presents a HGV-KM research model. This model identifies the context of HGV KM studies, the predispositions and factors that may impact KM outcomes, and important KM processes. It also represents the relationships among these issues. Applying the model, further studies will point the way for improved capture and dissemination of HGV knowledge from routine HGV research activities to contribute to the global genetics knowledgebase.

Keywords:

Knowledge Management (KM); knowledge processing; human genetic variation (HGV) research; KM approach; KM framework; KM process; HGV-KM research model

Introduction

Knowledge Management (KM) theories

Knowledge is a fluid mix of framed experience, values, contextual information and expert insight; it includes both explicit knowledge that is transmittable in formal systematic language and tacit knowledge that has a personal quality and is hard to formalize and communicate [1-3]. Different views of knowledge lead to different perceptions of knowledge management (KM) [4, 5]. For instance, (i) if knowledge is an object or information access, KM should focus on building and managing knowledge stocks; (ii) if knowledge is a process, the focus of KM is knowledge flow and knowledge processing – knowledge creation, acquisition, codification, retention, storage/ retrieval, integration, coordination, transfer, sharing, distribution, application, valuation and use; (iii) if knowledge is viewed as an organizational capability, then KM centers on building core competencies, understanding the strategic advantage of know-how, and creating intellectual capital [6-8].

Human Genetic Variation (HGV) context

Human genetic variation (HGV) research aims to characterize the nature, distribution and evolution of genetic variations in humans and to study the relationship *between* genetic variation *and* environment in the origins *and* characteristics of human populations *and* causes, diagnoses, treatments and prevention of the disease; HGV scientists come from various domains, e.g. genomics, proteomics and clinical sciences [9]. The clinical and laboratory genetic services test and interpret the HGV of patients and/ or families [10]. There were over 500 HGV testing laboratories in the United States at 1997 [11]. The HGV testing is becoming a routine procedure in clinics and research, with at least 751 active laboratories and 936 clinical chemistry/ hematology centers at 2004 in the European Union alone [12]. An international survey at 2005 shows that 45% of early breast cancer patients discuss genetic testing with their physician and/or are referred to see a genetic counselor and 16.7% are then tested [13]. However, despite the significant growth of HGV knowledge, KM success is seldom reported:

i) for managing high-quality *knowledge stocks*. Enormous efforts are put into storing HGV data in Locus Specific Databases (LSDB) and general databases, such as HGVbase [14], UMD LSDB [15], OMIM Database [16] and the proposed Central Database plus WayStation Submission tool [17-20]. However, the validity of data in these databases is of some concern, possibly due to inadequate data curation [21, 22].
ii) for supporting *knowledge flow* and enhancing *knowledge processes*. For example, although HGV testing laboratories collect and produce a lot of data with HGV details, knowledge flow from these laboratories to HGV research community is not occurring often [10, 23].
iii) for building core *competencies* and creating *intellectual capital*. With no attempt to improve the present "pattern" [24] of knowledge processing in HGV research facilities, the research ability and intellectual capital are not being managed or changed by KM efforts.

With above KM problems, we apply KM theories to identify all the relevant factors; then, we develop a HGV-KM research model to present the anatomy among the significant issues.

Implications of past literature

We synthesize past literature from leading Organizational Management journals (e.g., Organization Science and Harvard Business Review), Information Systems (IS)

periodicals, (e.g., MISQ and ISR), HGV journals (e.g., Human Mutation and Nucleic Acids Research), KM or HGV proceedings and books, then present the significant implications in this section.

KM approaches in practice

Most recorded KM practice took the *product-centric* or *process-centric* approach, reflecting type (i) or (ii) KM focus [6]; however, *capability-centric* KM exercise (type three) is rarely reported. The *product-centric* KM manage knowledge as an objective organizational asset [4, 25]. It relies on the transformation of implicit or explicit knowledge from employees' heads to written information in documents and the subsequent management of these documents [26]. Accordingly, by using searchable document repository and content management systems [7], HGV knowledge can be captured, stored, retrieved and distributed in well-organized research documentations.

On the other hand, *Process-centric* KM views knowledge as residing with a person and/or a business process. It provides pointers to experts [25] and implements business process management [27], by adopting database of experts, decision aids and expert system, workflow management system, groupware, the systems supporting 'Community of Practice' and 'hardwiring' of social networks, etc. [6, 28]. HGV studies require profound knowledge on the subjects and methodologies. Therefore, research done around the globe is frequently referred; and international collaborations are often performed. Taking these natures of HGV research into account, the key to managing HGV knowledge is to share it among the researchers [26], applying process-centric KM. Meanwhile, such KM approach may trigger the benchmarking, reengineering and optimizing of the HGV research processes. By tracking and sharing HGV research activities, instance decision making in single variant interpretation may be transited into best practice in studying the gene. This transition may eventually enhance an HGV research methodology and become a valuable intellectual capital, since medical data analysis may discover new models, into which available knowledge could be incorporated [29]. Thus, approaches in managing HGV knowledge may take all product-, process-, and capability-centric endeavors.

KM frameworks in literature

Following KM approach review, this subsection extracts the implications of KM models as they offer best KM practices.

The result of knowledge creation, retention and transfer is affected by the properties of organization units, of unit relationships and of knowledge itself; and this effect is moderated by three key causal mechanisms – the three important KM processes: ability, motivation and opportunity [8].

Based on organizational capability perspective theory [30] and contingency perspective theory [31], two more preconditions for effective KM are discovered – the 'knowledge infrastructure capability' and 'knowledge process capability' – with the latter being influenced by contingent knowledge tasks [32].

Knowledge management systems (KMS) are a class of IS to manage organizational knowledge, and to enhance knowledge processes [6, 33]. Developed upon IS Success Model [34, 35], Figure 1 shows how the individual's and organization's performance at workplace are improved from using KMS [36-38].

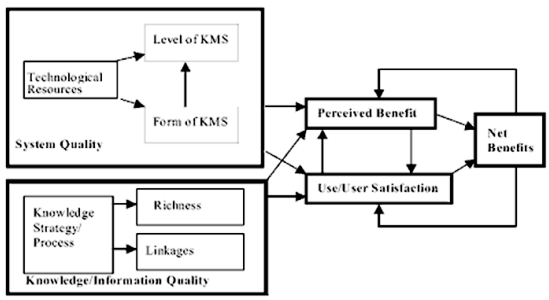


Figure 1 – KMS Success Model [36]

HGV-KM research model

To manage knowledge in HGV context, the practical questions are "which activities are the promising targets for KM support?" and "what are the nature and benefits of effective KM support?" An HGV KM study is to point the way for improved knowledge capture and dissemination from routine research activities to contribute to the global genetics knowledgebase. Drawing on past literature, we develop a research model (Figure 2) with nine significant KM issues: I. KM context, II. KM process, III. knowledge process capability, IV. contingent task characteristics, V. technology and system quality, VI. knowledge and information quality, VII. perceived benefits and user satisfaction, VIII. knowledge infrastructure capability, and IX. KM outcome. The construct relationships are direct impacts (as arrows from cause to result), moderating effects on a relationship's strength (by broken arrows) and weak connections (e.g. the quality of KM technologies as the extent to which knowledge processes and KM processes are computerized and integrated [36], represented by curves in the figure). In detail:

I. Context of a KM initiative or project [8]

According to [8], the predispositions of a HGV KM project include (i) the properties of units, e.g. a researcher's academic status and an organization's social status; (ii) the properties of relationships between units, e.g. the contact frequency between dyadic units and the connection pattern among multiple units; and (iii) properties of knowledge, including explicit HGV data and tacit know-how in HGV research, external and internal knowledge, and uniquely possessed and public knowledge.

II. KM process [8]

A KM process presents what a KM project offers and how well it functions to enhance an organization's (i) KM abil-

ity to codify implicit knowledge and produce information that makes sense to people other than the author [28], (ii) knowledge sharing motivation – the rewards and incentives, (iii) opportunity, and (iv) from process-centric KM perspectives, the management of HGV research activities and processes.

III. Knowledge process capability [32]

This organizational competency is the result of contextual factors and efforts (Constructs I and II) [8]; it then decides the quality of processed knowledge (Construct VI), according to product-centric KM theories and KMS Success Model [36]. It also has direct impact on overall KM effectiveness and this impact is moderated by knowledge-related tasks [32].

IV. Contingent knowledge tasks' characteristics [32]

The characteristics of knowledge tasks (such as task content and task domain) decide if the right knowledge is captured and used [38]. An example of task domain is about knowledge creation tasks that may belong to any of the four modes of socialization, externalization, combination and internalization [39]. The moderating role of this construct suggests that the KM efforts that precisely suit a task provide more effective results than those that don't [31, 32].

V. Technology and system quality [36]

In addition to the quality of general IS, such as 'ease of learning', 'integration of systems', and quick 'system response time' [34], KMS quality has three more dimensions: (i) the technological resources – the ability to develop, operate, and maintain a KMS, (ii) KMS form – the extent to which organizational memory and KM processes are computerized and integrated, and (iii) KMS level – the ability to bring past information to bear upon current activities [36, 37]. Given the teamwork nature of HGV research, quality KM technology should also support an HGV research facility's social capital (Construct VIII) by facilitating collaborations, distributed learning, knowledge mapping and opportunity generation [30, 40, 41].

VI. Knowledge and information quality [36]

High-quality knowledge and information are complete, accurate, current (of linkages), informative, rich in expression and in detail [34, 37]. For example, a valid HGV result has to offer reliable and sufficient evidence for variant interpretation based on accurate and unambiguous variant description.

VII. Perceived benefits and user satisfaction [36]

Perceptions on the benefits of KM technologies include perceived usefulness and ease-of-use, both of which are significant predictors of technology acceptance – the actual levels of system usage and user satisfaction [42]. This perspective of users is a result of knowledge quality; and it delivers ultimate KM outcomes in the organization [36, 38].

VIII. Knowledge infrastructure capability [32]

This capability represents the organization's social capital – the network of relationship; and it delivers KM results through knowledge sharing via the network [32]. It is operationalized by (i) technologies (Construct V), (ii) the organizational structure that provides the relationships (i.e. is an organization's property within Construct I) and (iii) the culture that provides a shared context (as a relationship's property in Construct I).

IX. Ultimate KM outcomes [36]

KM outcomes include the KM project-improved organizational effectiveness – such as the ability to innovate and coordinate [30] – and individual/organizational performance [38]. The individual KM performance, as measured by 'correctness of decision' and 'confidence in decision', will in turn have an impact on the organization's performance, e.g. on product quality [34, 36]. As a KM project's outcomes, the KM efforts may cause positive or negative consequences that will trigger more or less use of the knowledge and the KMS [38].

Discussion

From past KM approaches and frameworks that recorded the nature of knowledge, KM scope, KM factors and mechanisms, we identify nine categories of contextual,

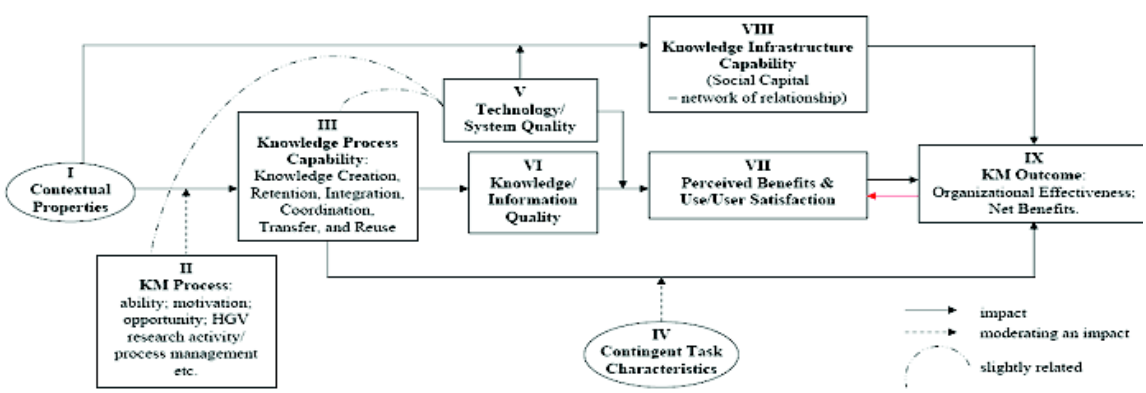


Figure 2 – HGV-KM Research model

cultural, structural, managerial, cognitive and technological issues that might be critical to the genetics research domain. Then we develop an HGV-KM research model presenting the dynamics among these nine constructs. There are a number of KM models proposed by both academics and practitioners that are not included in this paper, e.g. [43], [27], [24]. The reason for not covering them is that they focus more on single KM approach (product/process/capability-centric), but the HGV setting may have to apply all three strategies. Another limitation of our model is that it still needs validation, for instance, by empirical testing. However, it has revealed a promising research area that is to seek resolutions for the KM issues in genetics.

Our current research direction is to validate the model proposed herein through iterative action research. We plan to evaluate the impact of IT on various KM processes and capabilities, as well as on overall KM outcomes. This will also validate the model constructs and the model dynamics in terms of construct relationships; for instance, we are expecting to discover significant correlations between the features of a KM project (including implementation of a KMS) and resulting changes in KM performance. In addition to triggering more or less use of the knowledge and the KMS, feedbacks from KM effectiveness (Construct IX) might contribute to subsequent increases of user satisfaction, and possible establishments of KM-related organizational capabilities, and even changes in the contextual properties, such as the contact frequency and connection pattern in the social network. In conclusion, longitudinal studies on HGV KM practice may further refine our research model and add more insights on its anatomy.

Conclusion

Organizational knowledge management (KM) aims at effectively building and managing knowledge stocks, supporting knowledge flow and knowledge processes, building core competencies and creating intellectual capital. In the human genetic variation (HGV) research domain, KM efforts might improve the capture and dissemination of knowledge from routine HGV research activities to contribute to the global genetics knowledgebase. By synthesizing past literature, we have developed a KM research model with nine significant constructs; and we hope this paper will cast a light on future knowledge management research in the genetics domain.

Acknowledgments

We would like to thank many people for their support to this PhD project in the University of Auckland (UoA), especially the inspirational co-supervisor – Dr. Alexei Drummond, and our dearest colleagues Ms Karen Day, Dr. Martin Orr and Dr. Rekha Gaikwad. We are also indebted to Dr. Graeme Suthers, Prof. Andrew Shelling, Ms Deborah Prendergast, Ms Clare Stones and all the research students in the UoA Medical Genetics Group for their precious information and participation.

References

[1] Davenport TH, and Prusak L. Working Knowledge: How Organizations Manage What They Know, Cambridge, MA: Harvard Business School Press, 1997.

[2] Polanyi M. The Tacit Dimension, London: Routledge & Kegan Paul, 1966.

[3] Nonaka I. A dynamic theory of organizational knowledge creation. Organization Science 1994; 5 (1): 14-37.

[4] Schultze U, and Leidner DE. Studying knowledge management in information systems research: discourses and theoretical assumptions. MIS Quarterly 2002; 26 (3): 213-242.

[5] Carlsson SA, El Sawy OA, Eriksson I, and Raven A. Gaining competitive advantage through shared knowledge creation: in search of a new design theory for strategic information systems. In Proceedings of the Fourth European Conference on Information Systems, Lisbon, 1996; pp. 1067-1075.

[6] Alavi M, and Leidner DE. Review: knowledge management and knowledge management systems: conceptual foundations and research issues. MIS Quarterly 2001; 25 (1): 107-136.

[7] Sambamurthy V, Bharadwaj A, and Grover V. Shaping agility through digital options: reconceptualizing the role of information technology in firms. MIS Quarterly 2003; 27 (2): 237-263.

[8] Argote, L, McEvily B, Reagans R. Managing knowledge in organizations: An integrative framework and review of emerging themes. Management Sci 2003; 49(4 Special Issue) 571–582.

[9] The Human Genome Organisation (HUGO). Mission Statement. 2006; http://www.hugo-international.org/mission.htm, accessed at 2006-11-20.

[10] ITPerspectives Inc. NHS genetic service information systems output based specification. 2005; http://www.ngrl.org.uk/Manchester/Pages/Downloads/ITP-GEN3-OBSv05.pdf, accessed at 2005-11-03.

[11] Holtzman NA, Watson MS, eds. Promoting safe and effective genetic testing in the United States. Bethesda (MD): National Institutes of Health; 1997; http://www.genome.gov/10001733, accessed 2006-12-3.

[12] Ibarreta D, Elles R, Cassiman JJ, Rodriguez-Cerezo E, Dequeker E. Towards quality assurance and harmonization of genetic testing services in the European Union. Nat Biotechnol 2004; 22:1230-1235.

[13] Brown KL, Hutchison R, Zinberg RE, and McGovern MM. Referral and experience with genetic testing among women with early onset breast cancer. Genetic Testing Dec 2005; 9 (4): 301-305.

[14] The Karolinska Institute (Sweden) and the European Bioinformatics Institute (UK). Human Genome Variation Database. 2004; http://hgvbase.cgb.ki.se/, accessed at 2005-11-18.

[15] Laboratory of Human Genetics Montpellier (France). The Universal Mutation Database (UMD) Locus Specific Databases. 2005; http://www.umd.be/LSDB.html, accessed at 2005-11-18.

[16] National Center for Biotechnology Information (NCBI). OMIM – Online Mendelian Inheritance in Man. 2005; http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM, accessed at 2005-11-18.

[17] Horaitis O, and Cotton RGH. The challenge of collecting mutations across the genome: The Human Genome Variation Society approach, Hum Mutat 2004; 23: 447-452.

[18] Scriver CR, Nowacki PM, and Lehväslaiho H. Guidelines and recommendations for content, structure and deployment of mutation databases. Hum Mutat 1999; 13: 344-350.
[19] Scriver CR, Nowacki PM, and Lehväslaiho H. Guidelines and recommendations for content, structure, and deployment of mutation databases II: Journey in progress. Hum Mutat 2000; 15: 13-15.
[20] Genomic Disorders Research Centre. Welcome to the Waystation. 2006; http://www.centralmutations.org, accessed at 2006-11-20.
[21] Gu Y, Stanek J, and Warren J. Knowledge management (KM) technologies in medical genetic testing (MGT) laboratories – a literature review. In Proceedings of Health Informatics Conference (HIC 2006), Sydney, Australia, August 20-22, 2006.
[22] Fredman D, Munns G, Rios D, Sjöholm F, Siegfried M, Lenhard B, Lehväslaiho H, and Brookes AJ. HGVbase: a curated resource describing human DNA variation and phenotype relationships. Nucleic Acids Research 2004; 32(Database-Issue): 516-519.
[23] Gu Y, Warren J, Stanek J, and Suthers G. A system architecture design for knowledge management (KM) in medical genetic testing (MGT) laboratories. In Proceedings of the 10th International Conference on Computer Supported Cooperative Work in Design (CSCWD 2006), Nanjing, China, May 3-6, 2006.
[24] Firestone JM, and McElroy MW. Doing knowledge management. The Learning Organization 2005; 12 (2): 189-212.
[25] Dennis AR, and Vessey I. Three knowledge management strategies: knowledge hierarchies, knowledge markets and knowledge communities. MIS Quarterly Executive 2005; 4 (4).
[26] Bossen C, and Dalsgaard P. Conceptualization and appropriation: the evolving use of collaborative knowledge management. In Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility CC '05, August 2005, pp. 99-108.
[27] Massey AP, Montoya-Weiss MM, and O'Driscoll TM. Knowledge management in pursuit of performance: insights from Nortel Networks. MIS Quarterly 2002; 26 (3): 269-289.
[28] Brown JS, and Duguid P. Organizing knowledge. California Management Review 1998; 40 (3): 90-111.
[29] Zupan B, Holmes JH, and Bellazzi R. Knowledge-based data analysis and interpretation. Artificial Intelligence in Medicine 2006; 37 (3): 163-165.
[30] Gold AH, Malhotra A, and Segars AH. Knowledge management: an organizational capabilities perspective. Journal of Management Information Systems 2001; 18 (1): 185-214.
[31] Becerra-Fernandez I, and Sabherwal R. Organizational knowledge management: a contingency perspective. Journal of Management Information Systems 2001; 18 (1): 23-55.
[32] Lindsey, K., "Measuring Knowledge Management Effectiveness: A Task-Contingent Organizational Capabilities Perspective," Eighth Americas Conference on Information Systems, 2002, pp. 2085-2090.
[33] Ruggles RL. Knowledge tools: using technology to manage knowledge better. Cambridge, MA: Cap Gemini Ernst & Young, Center for Business Innovation, 1997; http://www.cbi.cgey.com/pub/docs/KnowledgeTools.pdf, accessed at 2001-09-01.
[34] DeLone WH, and McLean ER. Information systems success: the quest for the dependent variable. Information Systems Research 1992; 3 (1): 60-95.
[35] DeLone WH, and McLean ER. Information Systems Success Revisited. In Proceedings of HICSS, 2002, pp. 238-248.
[36] Jennex ME, and Olfman L. A knowledge management success model: an extension of DeLone and McLean's IS success model. In Proceedings of Ninth Americas Conference on Information Systems, August 2003.
[37] Jennex ME, and Olfman L. Organizational memory/knowledge effects on productivity, a longitudinal study. In Proceedings of 35th Hawaii International Conference on System Sciences, HICSS35, IEEE Computer Society, January 2002.
[38] Jennex ME, and Olfman L. Assessing knowledge management success/effectiveness models. In Proceedings of the 37th Hawaii International Conference on System Sciences, 2004.
[39] Nonaka I. The knowledge-creating company. Harvard Business Review 1991; 96-104.
[40] Grant R. Toward a knowledge based theory of the firm. Strategic Management Journal, Winter 1996; 17: 109-122.
[41] Leonard D. Wellsprings of Knowledge: Building and Sustaining the Source of Innovation. Boston: Harvard Business School Press, 1995.
[42] Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of infoamtion technology. MIS Quarterly 1989; 13 (3): 319-340.
[43] Bots PWG, and de Bruiin H, Effective Knowledge Management in Professional Organizations: Going by the rules. In Proceedings of the 35th Hawaii International Conference on System Sciences, 2002; IEEE Computer Society Press

Address for correspondence

Yulong Gu,
Department of Computer Science (Tamaki),
University of Auckland,
Private Bag 92019, Auckland,
New Zealand.
ygu@cs.auckland.ac.nz

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

ECTracker – An Efficient Algorithm for Haplotype Analysis and Classification

[1]Li Lin, [2]Limsoon Wong, [3]Tze-Yun Leong, [4]Pohsan Lai

[1;2;3] School of Computing, National University of Singapore
[4]Dept of Pediatrics, National University Hospital, National University of Singapore

Abstract

This work aims at discovering the genetic variations of hemophilia A patients through examining the combination of molecular haplotypes present in hemophilia A and normal local populations using data mining methods. Data mining methods that are capable of extracting understandable and expressive patterns and also capable of making predictions based on inferences made on the patterns were explored in this work. An algorithm known as ECTracker is proposed and its performance compared with some common data mining methods such as artificial neural network, support vector machine, naive Bayesian, and decision tree (C4.5). Experimental studies and analyses show that ECTracker has comparatively good predictive accuracies in classification when compared to methods that can only perform classification. At the same time, ECTracker is also capable of producing easily comprehensible and expressive patterns for analytical purposes by experts.

Keywords:

datamining, classification, hemophilia A, genetic variations, haplotypes

Introduction

In this paper, we propose a new algorithm, called ECTracker[1], for pattern extraction and classification of a specific type of biological dataset known as haplotypes. A total of 47 patients affected by hemophilia A and 47 matched normal controls from Singapore were genotyped with a set of markers located on chromosome Xq28 which tags the hemophilia A disease gene. Hemophilia A is an X-linked recessive bleeding disorder that results from deficiency and/or abnormality of coagulation factor VIII (FVIII) [1]. The FVIII gene spans 186 kb of DNA and resides on 0.1% of the X chromosome (band Xq28).

We are interested in methods that are capable of performing the two tasks efficiently – first to extract expressive patterns for descriptive analysis, and second to perform classification. Intuitively, expressive haplotype patterns (or genetic variations) need to be extracted to provide medical practitioners with insights about the genetic manifestations of patients affected by hemophilia A. The extracted patterns are used for predictive inference (or classification) to help in carrier detection, which is useful for medical prognosis and decision making.

In this paper, we present the design and implementation of the ECTracker method. We also examine its performance as compared to common data mining methods in supporting the targeted tasks. Specifically, we compared the expressiveness of the haplotype patterns discovered using ECTracker with the haplotype patterns discovered using the Decision Tree method (C4.5). Furthermore, we also compared the classification predictive accuracy of ECTracker with existing classification methods including Artificial Neural Network, Naïve Bayesian Network, Support Vector Machine and Decision Tree (C4.5) [2][3].

The ECTracker method

There are two main steps in ECTracker. First, it identifies the genetic variations (or haplotype patterns) of hemophilia A patients to help analyze FVIII gene polymorphism for linkage analysis. Second, the haplotype patterns found in the first step are used to perform classification to facilitate carrier screening by medical practitioners. Details of the hemophilia A dataset will be introduced in the next section.

Step 1 – Finding interesting patterns

The first step of the ECTracker algorithm uses a level-wise neighborhood search method to enumerate all possible marker patterns of length one, two, and three etc, and then computes the statistical *odds ratio* of each of the patterns. Only those patterns that are significant are selected. The significance of a potential/candidate pattern is determined by computing its p-value. P-value calculates the probability due to chance alone of getting a difference larger than or equal to that actually observed in the data [4] [5]. A small p-value means it is difficult to attribute the observed difference to chance alone, and this can be taken as evidence against the null hypothesis of non-significance.

Odds ratio is a test statistic that has been widely used in the biomedical arena to measure the magnitude of association between two categorical variables based on some data collected [6] [7]. Given a pattern x, odds ratio computes the ratio of non-association between x and the label L, to the association between x and L based on a set of data. For example, given a pattern, say (1,3), and there are σ number of such pattern found in a dataset D associated with the class label *Abnormal* and π number of such pattern found in D associated with class label *Normal*. We are interested

1 Initial findings of this work were presented as a poster in *Asia-Pacific Conference on Human Genetics,* 2004.

in finding out whether the marker pattern (1,3) is strongly associated with the label *abnormal*. Table 1 shows the contingency table for our example where *P* is the number of samples in the dataset associated with the class label *Abnormal* and *N* is the number of samples in the dataset with class label *Normal*. The odds ratio is computed based on equation 1 defined below.

Table 1 – 2x2 contingency table

	Abnormal	**Normal**
not(*1,3*)	$P - \sigma$	$N - \pi$
(*1,3*)	σ	π

$$\text{Odds Ratio, } \theta = \frac{(P-\sigma)\pi}{(N-\pi)\sigma} \qquad (1)$$

Step 2 – Predictive inference / classification

The following describes the algorithm for predictive inference using the patterns derived from the previous step. Before presenting the algorithm, let us define the order of precedence of the derived patterns. This is used in selecting patterns for our classifier.

Definition: Given two patterns, r_i and r_j, $r_i >> r_j$ (also called r_i precedes r_j or r_i has a higher precedence than r_j) if

1. The p-value of r_i is less than the p-value of r_j, the smaller the p-value of a pattern the greater the statistical significance of that pattern.
2. Both patterns have the same p-values and $r_i \subset r_j$, the pattern length of r_i is shorter than the length of r_j. The pattern with shorter pattern length that can correctly classify an unseen case is preferred.
3. Both patterns have the same p-values and $r_i \not\subset r_j$, but r_i is generated earlier than r_j.

Let *R* be the set of patterns derived in step 1, and *D* be the training data used to derive *R*. The basic idea of the algorithm is to choose a set of high precedence patterns in *R* as our classifier. The classifier is of the following format: $<r_1, r_2, \ldots, r_n, default\_class>$, where $r_i \in R$, $r_a >> r_b$ if $b > a$. The *default_class* is the chosen class for an unseen case when no pattern in the classifier could classify the unseen case. The *default_class* can be selected by the user. However, if the user decides to let the classifier select the *default_class*, then the majority class in the data *D* will be chosen as the *default_class*.

The algorithm for building the classifier consists of five steps:

$$\text{Total score for class } C_x, \ \Omega c_x = \sum \omega_{C_x} \qquad (3)$$

Step 1: Sort the set of generated patterns *R* according to the relation ">>". This is to ensure that we will choose the highest precedence patterns for our classifier.

Step 2: For each pattern *r* in sorted *R*, if there exist another pattern *r*' such that the p-values of both *r* and *r*' are the same, and $r' \subset r$, then remove *r* from sorted *R*. This ensures that we choose the pattern with the shortest pattern length for each p-value.

Step 3: Select the first *n* patterns from sorted *R* following the sorted sequence to form the set $\Re$ for classification.

Step 4: Perform classification on the training data *D* using the *n* pattern classifier $\Re$ and compute the true positive rate of the prediction.

Step 5: If the true positive rate is less than the user defined minimum true positive rate, then repeat Step 3 and Step 4 with a different *n* value.

In classifying an unseen case in Step 4, the first pattern that satisfies the case will classify it. If no pattern applies to the case, a scoring method will be used for each of the classes, where the class with the highest score classifies the case. However, if the scoring method produces the same score for each of the available classes, then the unseen case will take on the default class. The user is able to set the default class to "unknown" to allow the classifier to make no prediction when no pattern applies to the case. This is useful when there are samples that are identical in attribute values but belonging to different classes. Figure 1 shows the pseudocode for scoring the classes.

1. for each class C_x do
2. $Score(C_x) = 0$
3. for each pattern $r_i \in \Re$ do
4. if $r_i.class == C_x$
5. compute ω_{C_x}
6. $Score(C_x) = Score(C_x) + \omega_{C_x}$
7. end
8. end
9. end

Figure 1 - Pseudocode for computing score of each class

For each pattern r_i in $\Re$ that classifies a class C_x, computes the individual pattern score using equation 2 as follows:

$$\text{Individual pattern score of } r_i \text{ for class } C_x, \ \omega_{C_x} = \frac{(nummatch)^2}{patternlength - casepatternlength} \qquad (2)$$

Where *patternlength* refers to the pattern length of r_i and *casepatternlength* refers to the pattern length of the case to be classified, and *nummatch* refers to the number of attribute matches between r_i and the case pattern. The total score for a class C_x is computed as shown in equation 3 as follows:

The unseen case will take on the class with the maximum Ω value. The objective of counting partial matches is for better noise handling.

We now describe how the scoring scheme handles noise with an example. Given that a pattern ABC is significant, its subset AB may or may not be significant since the odds ratio value is neither upward nor downward closed. However, if at a different odds ratio value, it is found that both ABC and AB are the shortest significant patterns for a class C1, this would mean that the attributes A and B are

important for determining the attributes for the class C1. Now, if we have some other patterns say ADE and DEF that are significant for another class C2. If we now have a case pattern to classify, say A, then A will have a higher score for class C1. This is the desired effect since the pattern ADE for class C2 may become significant due to noise, but it is less likely for AB and ABC to become significant due to noise.

The hemophilia dataset

A set of five common PCR-based polymorphisms located on chromosome Xq28 which tags the hemophilia A disease gene were collected and analyzed from 47 patients and 47 matched normal controls. The five polymorphisms collected are two microsatellite repeats in introns 13 and 22, and three Restriction Fragment Length Polymorphisms (RFLPs), namely BclI-intron 18, HindIII-intron 19, and XbaI-intron 22.The exact location of the markers are shown in Figure 2.

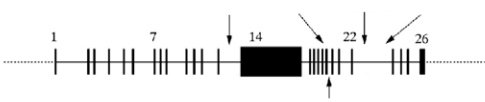


Figure 2 – Factor VIII gene

In the next sub-section, we describe the allelic frequencies of Factor VIII gene observed in our local population and the allelic frequencies reported by the authoritative resource website [8] for hemophilia A disease. The reporting of the allelic frequencies of our local population is useful for other medical practitioners not located in Singapore to decide whether they could make use of our discovery of the genetic variations for prognosis and counseling of their patients.

Allelic frequencies

The allelic frequencies observed in this study and those reported by *Hemophilia A Mutation, Structure, Test and Resources Site* [8] are tabulated in Tables 2, 3, and 4. Our results for *Bcl*I, *Hind*III, *Intron-13(CA)n* and *Intron22(GT)n(AG)n* are significantly similar to those reported in [35] with $\chi 2 < 3.841$ for *Bcl*I and *Hind*III, and $\chi 2 < 12.59$ for *Intron-13(CA)n* and *Intron22(GT)n(AG)n*. They are all within 95% confidence interval. However, the frequency for *Xba*I is significantly different from those reported by [35] with $\chi 2 > 3.841$.

Table 2 – Allelic frequencies of RFLPs

RFLPs	**Allele Frequencies (This Study)**		**Allele Frequencies (Reported by [35])**	
	(-)	(+)	(-)	(+)
	1	2	1	2
*Bcl*I	0.22	0.78	0.29	0.71
*Hind*III	0.78	0.22	0.75	0.25
*Xba*I	0.56	0.44	0.41	0.59

Table 3 – Allelic frequencies of Intron 13 (CA)n Repeats

	Allele Frequencies						
Intron 13 (CA)n Repeats	24	23	22	21	20	19	15
	1	**2**	**3**	**4**	**5**	**6**	**10**
This Study	0.01	0.10	0.06	0.26	0.52	0.04	0.01
Reported by [35]	0.013	0.05	0.11	0.29	0.45	0.07	0

Table 4 – Allelic frequencies of Intron 22 (GT)n/(AG)n repeats

	Allele Frequencies						
Intron 22 (GT)n/ (AG)n Repeats	31	30	29	28	27	26	25
	1	**2**	**3**	**4**	**5**	**6**	**7**
This Study	0.01	0.01	0.04	0.03	0.09	0.63	0.19
Reported by [35]	0	0	0	0.013	0	0.667	0.307

It was observed that samples with *Bcl*I-intron 18 allele 1 were always associated with *Hind*III-intron 19 allele 2 with $\chi 2$ p-value < 0.001. This is an expected observation as there are reported linkage disequilibrium between *Bcl*I and *Hind*III alleles from literature such as Ahrens et. al. [9] and EL-Maarri et. al. [10]. The *Hind*III marker was thus excluded since *Bcl*I and *Hind*III are in linkage disequilibrium, we could easily predict the value of the other attribute base on the value of one attribute, and hence 4 markers are sufficient in for the analysis.

It was further found that 70% of the samples had exactly the same allele values in all the markers in both patient and normal controls, which means that the 5 markers/attributes in the dataset is insufficient for separating 70% of the samples. After removing those samples whose disease and normal haplotypes cannot be distinguished, there are 28 samples remaining – 18 samples belonging to the disease phenotype and 10 samples belonging to the normal/control phenotype. Tables 5 and 6 show the frequencies of the disease and normal/control haplotypes respectively.

For descriptive analysis, we mainly report on the expressive and interesting patterns extracted from the remaining 30% of the dataset. Whereas for classification or predictive analysis, we divide the experiment into two parts: The first part compares the accuracies of the five classifiers based on the full hemophilia dataset. The second part we concentrate our study on the 30% of the dataset where those samples whose disease and normal haplotypes cannot be distinguished were removed.

Table 5 – Haplotype frequencies of probands with disease phenotype

Marker	Disease Haplotypes											Total
Intron-13 (CA)n	3	4	4	4	4	4	5	5		5	10	
*Bcl*I	1	2	2	2	2	1	2	2	2	1	2	
*Xba*I	1	1	1	2	2	1	2	1	2	1	1	
Intron-22 (GT)n/(AG)n	3	1	3	3	5	7	2	4	5	6	6	
No. of Probands	1	1	2	1	1	6	1	2	1	1	1	18

Table 6 – Haplotype frequencies of probands with normal/control phenotype

Marker	Normal/Control Haplotypes									Total
Intron-13 (CA)n	1	2	2	3	3	4	4	5	6	
*Bcl*I	1	1	2	1	1	2	2	2	2	
*Xba*I	1	1	2	1	1	1	1	1	2	
Intron-22 (GT)n/(AG)n	7	5	6	5	7	4	5	7	6	
No. of Probands	1	1	1	1	2	1	1	1	1	10

Results

Interesting pattern extraction

Expressive patterns derived by C4.5

C4.5 deduced that haplotype patterns (or genetic variations) of 4–*–*–*, 5–*–*–*, or 10–*–*–* (Intron13(CA)n–BclI–XbaI–Intron22(GT)n(AG)n) are highly associated with the disease phenotype. This derivation is not very useful as we could see from Table 6 that there are 3 probands with normal/control phenotype having intron-13 (CA)n allele values 4 and 5. Moreover, allele value 10 in intron-13 (CA)n only occurs once in the proband with disease phenotype (from Table 5), hence it is not able to give a generalize conclusion base only on allele value 10 of intron-13 (CA)n.

The possible reason for such a deduction by C4.5 may be due to the problem that the dataset is very small, and as a result the selection for partitioning attribute becomes biased for those attributes with more attribute values. Hence attributes with more attribute values will be assigned higher information gain as compared to attributes with fewer attribute values.

Expressive patterns derived by ECTracker

The longest most significant pattern associated with the disease phenotype derived by ECTracker is 4–1–1–7 (Intron13(CA)n–*Bcl*I–*Xba*I–Intron22(GT)n(AG)n). This is an interesting observation as the haplotype occurs in 33.3% of the disease phenotype and 0% of the normal/control phenotype with $\chi 2 > 3.841$ and odds ratio $\theta = 0$, which means that such observation occurs significantly greater than by chance. From Table 5, the haplotype occurs in 6 probands with disease phenotype as compare to other haplotypes which occur in no more than 2 probands. The shortest most significant patterns derived by ECTracker are 4–*–*–7 or 4–1–*–* with $\chi 2 > 3.841$ and odds ratio $\theta = 0$. This means that two markers alone were sufficient to define the disease haplotype, however, the longest most significant pattern provides a useful insight for the medical practitioners or scientists who seek to better understand the genetic variations of the disease.

This experiment shows that as compared to the decision tree approach of C4.5, ECTracker is capable of deriving useful patterns even when the dataset is very small.

Classification of the hemophilia A dataset

There are a total of 94 records in the hemophilia dataset, 47 records belonging to the class patient and 47 records belonging to the class normal. The classification methods that we examined include C4.5, Naïve Bayesian Classifier, Neural Network, Support Vector Machine and ECTracker. Except for ECTracker, all the other four classification algorithms are available from the package WEKA. WEKA is an open source data mining and machine learning software [2]. Table 7 shows the performance of various classifiers when applied to the full hemophilia dataset, and Table 8 shows the performance of the classifiers when applied to the pruned hemophilia dataset.

Table 7 – Analysis of classifiers based on full hemophilia dataset

	Accuracy	Precision of Class *Patient*	Recall for Class *Patient*	Precision for Class *Normal*	Recall for Class *Normal*
C4.5	71.43%	0.708	0.944	0.75	0.3
Naïve Bayesian Network	64.29%	0.7	0.778	0.5	0.4
Artificial Neural Network	78.57%	0.833	0.833	0.7	0.7
Artificial Neural Network	71.43%	0.75	0.833	0.625	0.5
Support Vector Machine	82.14%	0.842	0.889	0.778	0.7

Table 8 – Analysis of classifiers based on pruned hemophilia dataset

	Accuracy	Precision for Class *Patient*	Recall for Class *Patient*	Precision for Class *Normal*	Recall for Class *Normal*
C4.5	62.77%	0.615	0.681	0.643	0.574
Naïve Bayesian Network	62.77%	0.615	0.681	0.643	0.574
Artificial Neural Network	64.89%	0.675	0.574	0.63	0.723
Support Vector Machine	63.83%	0.623	0.702	0.659	0.574
ECTracker	67.02%	0.674	0.660	0.667	0.681

The classifiers were evaluated using 5-fold cross validation for the full dataset and leave-one-out for the pruned dataset. Since for each fold, different data samples are selected it is necessary run both steps in Section 2 fresh to avoid bias. Therefore, the patterns used for classification will vary for each fold.

ECTracker outperformed other classifiers with higher predictive accuracies on both full and pruned hemophilia A datasets.

On the full hemophilia dataset, the predictive accuracy of ECTracker is 67.02%, followed by Artificial Neural Network and Support Vector Machine with predictive accuracies of 64.89% and 63.83% respectively. Both Naïve Bayesian Network and C4.5 have the same predictive accuracy of 62.77%.

On the pruned dataset, ECTracker was able to accurately predict the phenotype of a sample given its polymorphic markers 82.14% of the time, followed by Artificial Neural Network at 78.57% of the time. C4.5 and Support Vector Machine were able to make accurate predictions 71.43% of the time.

There are 94 records in the unprune hemophilia A dataset, which is reasonably large, and there are 23 records in the pruned hemophilia A dataset which is rather small. Our experiments show that ECTracker is capable of providing good classification accuracy on both small and large datasets when compared to other classification methods.

Conclusion

In this work, we explored methods that are capable of extracting understandable and useful patterns, and also capable of performing inference on the patterns to make prediction. We applied these methods to find the genetic variations of a real dataset consisting of patients affected by hemophilia A to facilitate haplotype analysis by medical practitioners. We examined the issues of descriptive and predictive analyses using our proposed method called ECTracker.

In descriptive analysis, ECTracker is capable of extracting comprehensible and useful patterns from the hemophilia A dataset. Comparing with the patterns derived by C4.5, the patterns derived by C4.5 are less useful, as described earlier.

In predictive analysis or classification, ECTracker is capable of producing good predictive accuracies in classification that are comparable to those methods that only perform classification such as Artificial Neural Network and Support Vector Machine.

The experiments have indicated that ECTracker is potentially an effective method for both pattern extraction and classification for biomedicine in particular, and datamining in general.

The approach proposed here provides analysis and classification based on mainly the disease status of an individual. Continuously distributed quantitative traits such as blood pressure and choltesterol level may also be of significance to the clinicians. ECTracker can be extended to perform analysis and classification based on continuously distributed quantitative traits by defining a new scoring method for the interesting patterns. Further investigation will need to be done to assess the feasibility of such extension.

References

[1] Antonarakaris S, Kazazian H, Tuddenham E. Molecular etiology of factor VIII deficiency in hemophilia A. Human Mutation 1995; 5:1-22.

[2] Witten IH, Frank E. Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann; 2000.

[3] Michie D, Spiegelhalter D, Taylor C. Machine Learning, Neural and Statistical Classification. Ellis Horwood: New York; 1994.

[4] Ott L. An Introduction to Statistical Methods and Data Analysis. PWS-Kent Publishing Company, Massachussetts; 1988.

[5] Theodorou, S.J., Theodorou, D.J. and Kakitsubata, Y. Statistical Analysis in Clinical Studies: An Introduction to Fundamentals for Physicians. Internet Medical Journal, 2004.

[6] Le Moual N, Siroux V, Pin I, Kaufmann F, Kennedy SM. Asthma Severity and Exposure to Occupational Asthmagens. American Journal of Respiratory Critical Care Medicine 2005.

[7] Deheinzelin D, Lourenco MT, Costa CL, Younes RN. The Level of Nicotine Dependence is an Independent Risk Factor for Cancer: A Case Control Study. Clinics 2005; 60(3):221-226.

[8] HAMSTeRS *Haemophilia A Mutation, Structure, Test and Resource Site* http://europium.csc.mrc.ac.uk/WebPages/Main/main.htm.

[9] Ahrens P, Kruse TA, Schwartz M, Rasmussen PB, Din N.A. New HindIII Restriction Fragment Length Polymorphism in the Hemophilia A Locus. Human Genetics 1987; 76(2):127-128.

[10] EL-Maarri O, Kavakli K, Caglayan H. Intron 22 Inversions in the Turkish Haemophilia A Patients: Prevalence and Haplotype Analysis. Haemophilia 1999; 5(3):169-173.

Address for correspondence

Email: hi.linli@gmail.com

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

A Dynamic Query System for Supporting Phenotype Mining in Genetic Studies

Angelo Nuzzo, Daniele Segagni, Giuseppe Milani, Carla Rognoni, Riccardo Bellazzi

Department of Computer Science and Systems, University of Pavía, Pavía, Italy

Abstract

This paper describes an information technology infrastructure aimed at supporting translational bioinformatics studies that require joint management of phenotypic and genotypic data. In particular, we integrated an electronic medical record with an open-source environment for data mining to create a flexible and easy to use query system aimed at supporting the discovery of the most frequent complex traits. We propose a logical formalization to define the phenotypes of interest; this is translated into a graphical interface that allows the user to combine different conditions relative to the electronic medical record data (e.g., the presence of a particular pathology). The phenotypes are then stored in a multidimensional database. Then, the data mining system engine reads the filtered data from the database and executes dynamic queries for analyzing phenotypic data, presenting the results in a multidimensional format through a simple web interface. The system has been applied in a study on genetically isolated individuals, the Val Borbera project.

Keywords:

phenotype mining, complex traits, intelligent query, clinical data warehouse.

Introduction

A specific characteristic of the post-genomic era will be the correlation of genotypic and phenotypic information [1][2]; the emerging discipline of Biomedical Informatics may provide knowledge and tools for dealing with such an ambitious goal [3]. In this context, the studies aimed at the so-called genetic dissection of complex traits represent a first crucial benchmark for Biomedical Informatics and for translational bioinformatics.

The definition of an Information Technology infrastructure to support this kind of studies, and in particular the studies aimed at the analysis of large sets of phenotypes to discover the most prevalent diseases and then to integrate genotypic information, is a challenge which can be considered a paradigmatic goal of Biomedical Informatics. As a matter of fact, research on phenotypes requires the definition of an architecture for data collection, the implementation of an electronic medical record, the development of a system for the definition of the phenotypes of interest, and the design and implementation of a data warehouse system for analyzing phenotypic data. Moreover, selecting which are the phenotypes of interest determines the subsequent genotyping choices, especially when a genome-wide scan is not feasible or suitable.

Once clinical data are collected, it is crucial to perform a series of queries and data aggregation steps to characterize the population and extract the most prevalent phenotypes. However, clinicians, biologists and epidemiologists are usually unable to explore the collected information, because the use of general query languages requires substantial technical skill, as well as knowledge of the underlying database structures. On the other hand, the need of performing "dynamic" queries hampers the implementation of a "standard" user interface for pre-defined queries. To address this problem, we are defining a dynamic query system based on data warehouse and mining concepts. The clinical data are copied into a data mart oriented to data analysis. Thanks to the integration in the overall system of open-source environment for data mining it is possible to design a simple interface for performing aggregation, counting and simple statistics on the majority of the variables contained in the clinical database.

The implementation of the system has a number of steps which follows a workflow targeted at identifying the most important phenotypes which characterize a particular population. Such workflow includes the following steps:

1. The development of a relational database collecting clinical data on the target population
2. The translation of the database structure into a multidimensional data-base (data mart) oriented for query and reporting
3. The formal definition of the phenotypes to be searched and studied and their mapping in the database
4. Finally, the design and implementation of the data mining tool to easily extract the phenotypes and analyze their relationships

The aim of the paper is to describe the IT infrastructure of the system and the biomedical informatics challenges we are dealing with. Preliminary results on the use of the system to support a study on genetically isolated individuals, the Val Borbera project [4], will also be reported.

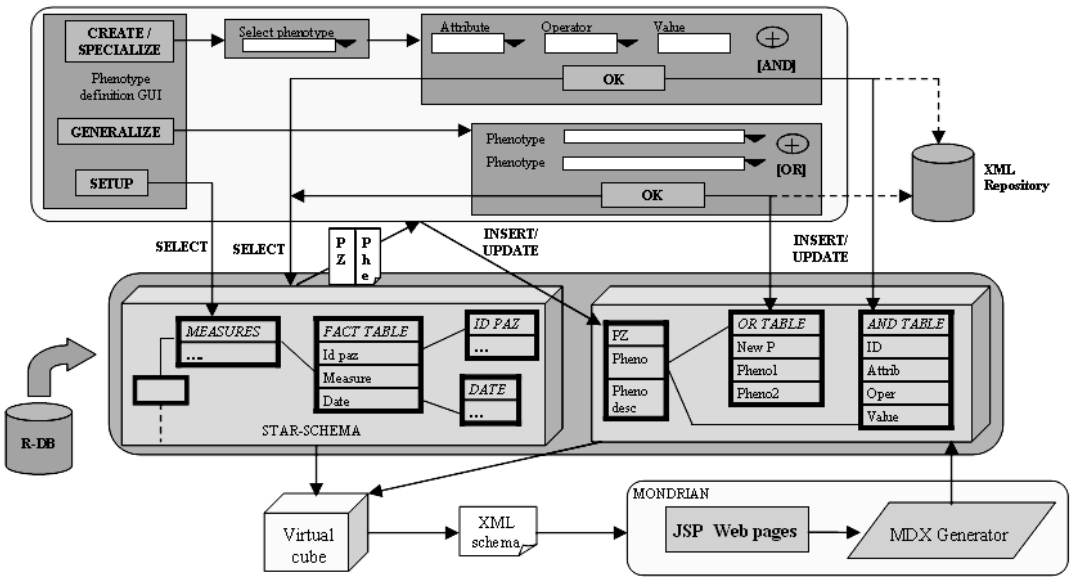


Figure 1 - Overview of system's components an the interactions among the 3 layers

Materials and methods

Objective

The final purpose of the system that we have developed is to provide a versatile and easy to use data inspection tool to identify which phenotypes may be more successfully investigated in the population under study, so that suitable genotyping choices may be made subsequently. The system accomplishes this purpose by providing tools for two main tasks:

1. for formally define the phenotypes to be investigated by a graphical user interface
2. for exploring clinical data to extract and analyze individuals with the same phenotypes (as they have been previously defined)

Both aspects can be performed in a database where the information is stored in a non-normalized data structure, which is often referred to as data mart. As a matter of fact, while a normalized structure is always required for correct management of the database (as it concerns data security, integrity, reliability), the database is greatly facilitated by the use of query-oriented data schema. That's why the basic common layer of our implemented system is a non-normalized data structure, the so-called "star schema", to which the two developed applications for query definition and data assessment are interfaced. Details of the components and their interactions are described in the following paragraphs, while a global overview is given in Figure 1.

From the relational DB to the "star-schema"

The typical data structure that has become a standard for all data warehouse applications is a multi-dimensional model called the "star join schema". Unlike the Entity-Relation model, the dimensional model is asymmetric. There is one large dominant table in the center of the schema, called the fact table. It is the only table in the schema which is connected to the other tables with multiple joins. Such other tables, called the dimensional tables, only require a single join to be referenced by the fact table [5].

Typically a clinical database can be modeled by a star schema in which each record in the fact table represents a combination of a clinical measure and its values on a specific date for a specific patient. So the dimensions are individuals, measurement time and measurement values: all of them can be further specified using a snowflake model[1].

The adaptation of the star-schema and of the snowflake models to the clinical context requires however several efforts. In fact, when taking into account the phenotype information in the analysis, we cannot model it as a dimension of the fact table, because it would be a *non-additive* dimension with respect to the others. *Additivity* is the ability to use an aggregate operator (summation, counting, average) along the dimensions of the same fact [6]: in our case, the phenotype dimension would be additive only along the patient dimension, but not along the others (measurements and time), as it is defined by a set of measurements. To overcome this problem, we have defined a new fact table to model the relationship between phenotypes and individuals.

The star schema and the phenotype tables are the physical models of the new multidimensional database. They represent two multidimensional "cubes" that together form the logical model of the database. The cubes may be merged in a single "virtual" cube, so that it is possible to use an Online Analytical Processing (OLAP) engine to perform

1 A *snowflake model* is a model in which a given dimension has relationships to other levels of the same dimension. It is used to re-normalize complex dimensions to eliminate redundancy.

data analysis (described in detail in the following paragraphs).

The phenotype definition tool

Clinicians and biologists usually define a phenotype by a set of variables and the values they may take. In order to select (and then to analyze) the individuals satisfying that set of rules, it is necessary to write a suitable SQL statement to run a query to retrieve them. However, as the users may have no expertise in the use of a query scripting language, we provide a tool that automatically generate the proper SQL script to select individuals with the defined phenotype.

To perform this task, a formalization of the phenotype definition is needed. The basic assumption is to consider a phenotype as a set of conditions in the form of attribute/value pairs. Then using logical operators (AND, OR) it is possible to combine different conditions to define more and more complex phenotypes. In particular, the AND operator allows the specialization of a defined phenotype, while the OR operator is used to merge different phenotypes into a single, more comprehensive one. This procedure corresponds to a logical tree construction, in which the nodes are the conditions, the AND operator is used to go from the top to the bottom and the OR operator is used to add an upper node from the bottom to the top (figure 2).

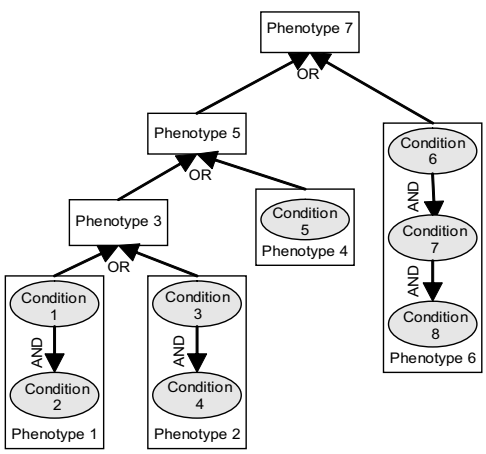


Figure 2 - Formalization of the phenotype definition

The limitation of this model is that it is not possible to define any kind of AND/OR combination of the conditions under investigation, as in the corresponding SQL string, it will not be possible to interpret the correct order of the subgroups (i.e., the correct position of brackets in the query string). So it is necessary to define the AND-conditions first, and then combining them using the OR operator.

The end users can create these definitions using a graphical wizard developed in the JAVA programming language. It interacts with the two sections of the non-normalized database, i.e., the star schema of the clinical data and the phenotype definition tables. The wizard is automatically setup by reading an XML file in which the two data structures are encoded. The conditions (attributes and their values) may be defined by combo boxes, which provide lists of attributes according to the measures table of the star schema, suggesting the admissible ranges of values for each attribute. Once the rules are defined, the corresponding SQL string is created by merging conditions by AND operators, in order to: i) store the rules in the phenotype section tables, and ii) select the subgroup of individuals satisfying that conditions and storing the relation individuals-phenotype. In the same way, it is possible to choose another graphical panel to selected some defined phenotypes to be merged together by the OR operators to create a new phenotype, which is stored in the phenotype section.

The data analysis tool

Dealing with clinical data to analyze phenotypic information implies taking into account heterogeneous data and viewing them at the same time. This means that it should be possible to perform a multidimensional inspection of the dataset. Whereas a relational database stores all data in the form of rows and columns, a multidimensional dataset consists of axes and cells organized in multidimensional "cubes", the dimensions of which are the directions of drill-down investigations.

The technique of multidimensional analysis is implemented in software tools called online analytical processing (OLAP) engines., OLAP, in fact, means analyzing large quantities of data in real-time. Unlike Online Transaction Processing (OLTP), where typical operations read and modify individual and small numbers of records, OLAP deals with data in bulk, and operations are generally read-only. The term "online" implies that even though huge quantities of data are involved — typically many millions of records, occupying several gigabytes — the system must respond to queries fast enough to allow an interactive exploration of the data.

As described above, the logical model of the star schema and the phenotype tables consist of two virtual cubes that may be merged in a single one, which is the input for the OLAP engine. In our system we use an OLAP engine written in the Java programming language: Mondrian [7]. It executes queries written in the MDX language (that has actually become a standard for data warehouse applications) [8], reads data from a relational database, and presents the results in a multidimensional format through a Java API, so that the presentation layer may be chosen by the final user. JSP pages are provided by default, so that the user can simply use a web browser for data visualization. The MDX queries have to be defined by the user, so we have developed a specific module (the "MDX generator" box in figure 1) that automatically create the MDX scripts directly from the attributes of interest chosen by the check box lists of the main page.

Sharing and generalizing issues

In order to make the graphical wizard usable on different databases, it has been made configurable via an XML file containing the star schema description. So the only prereq-

uisite for using it is to provide a star schema that is compliant to the model described above. Then the phenotype tables are automatically generated and populated by the GUI. Moreover, the phenotype definitions are also stored in XML files, so that existing phenotypes loading is performed by reading the XML files instead of the tables, and the XML repository may be shared with other scientists interested in analogue analysis. On the other side, whichever OLAP engine will be chosen, the only manual task needed is to code the virtual cube in the specific format required as input by the engine. Using Mondrian, it means to create the XML file containing the definition of the cube.

Results

The tools described in the previous section have been tested for the exploration of the clinical database of the Val Borbera genetically isolated population project [4]. This study is conducted in collaboration with the DIBIT of San Raffaele Scientific Institute of Milan, for which we have provided the architectural IT infrastructure for data collecting and storing. The clinical data have been collected in a relational database in a high normal form, actually containing about one hundred clinical measures relative to more than 4000 individuals. So the first step was to create the correspondent star-schema for the multidimensional analysis. Here we present an example regarding the analysis of dysfunctions related to the thyroid.

Phenotype definition

Before the use of a multidimensional analysis approach, the biologists had to ask a technician to extract individuals with the traits of interest by writing the SQL statement to be executed on the relational database. The statements are often some hundreds of lines long, due to the large number of join to be performed. Using the developed infrastructure, in contrast, the user has only to define the conditions by the graphical window shown in figure 3, that will be merged by the AND operators. The combo box are automatically filled in by reading the XML file in which the star schema is encoded.

Figure 3 - Creation/specialization panel

The defined phenotypes are summarized in the main panel of the GUI. The phenotype list shows the XML files that have already been created; details of the rules applied to define it are given below by selecting one of them (figure 5).

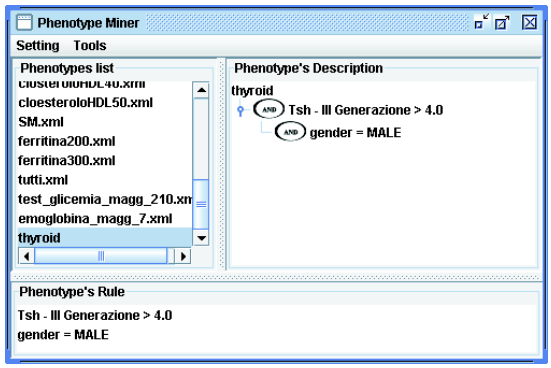


Figure 4 - Main panel: existing phenotypes are shown and details for the selected one

So if the user wants to specialize an existing phenotype, he can select it from the "Phenotype's Rules" window (figure 3) and add conditions using the AND operators.

Otherwise, if some defined phenotypes have to be merged together by OR operators to create a new phenotype definition, they can be selected using the "generalization panel", as shown in figure 4. In this case the combo box are filled by reading the XML phenotype files stored in the suitable repository.

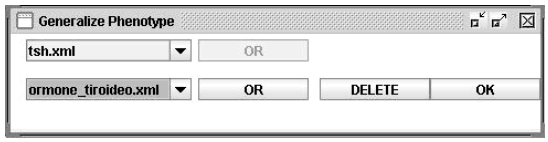
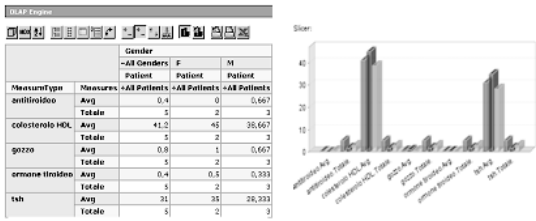


Figure 5 - The generalization panel

In both cases, the selection of individuals with the specified phenotypes is performed, and the relations are stored in the proper table. So the phenotype becomes a dimension which can be used to explore individual's data.

Data analysis and validation

When the OLAP engine starts up, it reads the XML file containing the cube definition. The first page shows a set of check boxes containing the fields of the underlying tables, so that the user may choose the variable to be investigated (the phenotypes are among them). Once the features have been chosen, the engine loads information related to the individuals having that phenotype. Then a visual inspection of the measurements values can be done expanding or collapsing cells of the resulting table, so that the analysis can be executed at different levels of detail (figure 6). Automatic graphical reports can also be generated.



		Gender		
		-All Genders	F	M
		Patient	Patient	Patient
MeasureType	Measures	+All Patients	+All Patients	+All Patients
antitiroideo	Avg	0,4	0	0,667
	Totale	5	2	3
colesterolo HDL	Avg	41,2	45	38,667
	Totale	5	2	3
gozzo	Avg	0,8	1	0,667
	Totale	5	2	3
ormone tiroideo	Avg	0,4	0,5	0,333
	Totale	5	2	3
tsh	Avg	31	35	28,333
	Totale	5	2	3

Figure 6 - The resulting dynamic table provided by Mondrian: the user may expand or collapse the cells to view the data at different levels of detail

A first validation of the system is now undergoing; it consists of verifying how it improves the phenotype searching process. Biologists formulate several verbal requests, we then compared the results obtained by themselves using the system and the ones obtained by a technician performing a SQL statement (which is typically quite a lot of instruction lines long). The first initial report table is always the same for both tasks and for any condition exploited (many different phenotypes corning thyroid diseases, hypertension and diabetes have been defined and searched): this means that the query editor correctly runs. The advantage provided by the OLAP engine is then the capability of allowing a dynamic interaction with the data to perform a more detailed exploration.

Conclusion

Current genetic studies are characterized by the collection of huge quantity of both clinical and genotypic data. The final goal of such effort is the "genetic dissection" of complex phenotypes; the first challenge of data analysis is therefore to identify which phenotypes must be investigated. However, this task may be difficult to be performed using "standard" tools for database navigation, such as SQL query, as they require technical skills for the end user to extract interesting information. In order to solve this problem, in this paper we present a dynamic query system based on data warehouse and mining concepts, which allows phenotypes definition by a graphical user interface and data exploration using OLAP tools. Both applications are based on a common underlying data layer, the structure of which is a data mart oriented to data analysis. The phenotype definition is based on a logical formalization and it is properly stored to be processed by the OLAP engine. Once the phenotypes have been defined, the OLAP engine allows the user to perform a visual inspection of the data through a set of results dynamically created. We have chosen an open source OLAP engine, Mondrian, and we integrated it with new components (the MDX generator) in order to automate other specific technical operations.

All the modules of the system are configurable via XML files, so that they can be reused to analyze other clinical databases. The only requirement is to translate the data structure into the data mart described in this paper, and to codify it in an XML file. Moreover, the phenotype definitions are also stored in an XML repository, so they may be reused and shared with other users to compare results.

The system as been tested on a real dataset, the clinical database of the Val Borbera project, showing that it is easy to use and time-saving. Future development of the system will improve the graphical user interface. The phenotype definitions will be shown by graphs, corresponding to the logical model used to create it, so that the user may expand it directly by adding or removing nodes.

Acknowledgments

We gratefully acknowledge Cristiana Larizza for software development supervision and DIBIT group (D. Toniolo, C. Sala, C. Camaschella) for providing data of the Val Borbera project.

References

[1] Lander ES, Schork NJ. Genetic dissection of complex traits. Science. 1994 Sep 30; 265(5181):2037-48.

[2] Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. Nature Genetics, 2003 Mar; 33:228-237.

[3] Martin-Sanchez F, Iakovidis I, Norager S, Maojo V, de Groen P, Van der Lei J, Jones T, Abraham-Fuchs K, Apweiler R, Babic A, Baud R, Breton V, Cinquin P, Doupi P, Dugas M, Eils R, Engelbrecht R, Ghazal P, Jehenson P, Kulikowski C, Lampe K, De Moor G, Orphanoudakis S, Rossing N, Sarachan B, Sousa A, Spekowius G, Thireos G, Zahlmann G, Zvarova J, Hermosilla I, Vicente FJ. Synergy between medical informatics and bioinformatics: facilitating genomic medicine for future health care. JBI, 2004 Feb; 37(1):30-42.

[4] Sala C, Bione S, Crocco L, Gatti M, Poggiali E,. Bellazzi R, Buetti I, Rognoni C, Camaschella C and Toniolo D. The Val Borbera Project: epidemiological and genealogical analysis of an isolated population in Northern Italy. Submitted to European Society of Human Genetics, 2006.

[5] Kinball R, Ross M. The Data Warehouse Toolkit, Second Edition. Wiley and Sons, Inc., 2002.

[6] Wyderka KC. Data Warehouse Technique for Outcomes Management. Health Management Technology, Nov. 1999; 20 (10):16-7.

[7] Hyde J, Mondrian OLAP project, Pentaho Analysis Service. http://mondrian.pentaho.org/.

[8] Spofford G, Harinath S, Webb C, Huang DH, Civardi.F. MDX Solutions, Second Edition. Wiley Publishing, Inc., 2006

Address for correspondence

angelo.nuzzo@unipv.it

This page intentionally left blank

Chapter 10.

Biomedical Image and Signal Processing

This page intentionally left blank

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Identifying QT prolongation from ECG impressions using Natural Language Processing and Negation Detection

Joshua C. Denny, MD[a,b], Josh F. Peterson, MD[a,b]

Departments of [a]Biomedical Informatics and [b]Medicine, Vanderbilt University Medical Center, Nashville, TN, USA

Abstract

Electrocardiogram (ECG) impressions provide significant information for decision support and clinical research. We investigated the presence of QT prolongation, an important risk factor for sudden cardiac death, compared to the automated calculation of corrected QT (QTc) by ECG machines. We integrated a negation tagging algorithm into the KnowledgeMap concept identifier (KMCI), then applied it to impressions from 44,080 ECGs to identify Unified Medical Language System concepts. We compared the instances of QT prolongation identified by KMCI to the calculated QTc. The algorithm for negation detection had a recall of 0.973 and precision of 0.982 over 10,490 concepts. A concept query for QT prolongation matched 2,364 ECGs with precision of 1.00. The positive predictive value of the common QTc cutoffs was 6-21%. ECGs not identified by KMCI as prolonged but with QTc>450ms revealed potential causes of miscalculated QTc intervals in 96% of the cases; no definite concept query false negatives were detected. We conclude that a natural language processing system can effectively identify QT prolongation and other cardiac diagnoses from ECG impressions for potential decision support and clinical research.

Keywords:

electrocardiogram, Unified Medical Language System, natural language processing, concept identification, decision support.

Introduction

Electrocardiograms (ECG) provide significant medical information that has been largely untapped in medication decision support interventions. ECGs are commonly used to help diagnose cardiac diseases such as myocardial infarction and ischemia, arrhythmias, and cardiomyopathies as well as some extracardiac diagnoses. Recent studies have highlighted the risk of sudden cardiac death due to medications known to increase QT intervals [1, 2]. Medications can also affect the risk of myocardial infarction [3, 4], induce second and third degree atrioventricular block [5], and cause other, potentially fatal, arrhythmias [1]. A decision support system that warns providers when prescribing medications to patients with existing risk factors, such as QT prolongation or atrioventricular block, may be valuable in guiding prescription choice.

QT prolongation is a key risk factor for development of Torsades de Pointes, a potentially fatal cardiac dysrhythmia. Since the QT interval, the measurement of the time between ventricular contraction ("QRS complex") and its repolarization ("T wave"), varies with heart rate, QT prolongation is typically assessed via a QT corrected for rate (QTc) [6]. A value greater than 440-460 ms is typically considered prolonged. Many drugs are known to prolong the QT interval; it is the most common reason for a drug to be removed from the market [2, 7]. However, many other factors can influence measurement of the QT (and thus QTc), including arrhythmias, intraventricular conduction disturbances, ECG measurement technique, and morphological changes in the ECG.

ECG findings consist of two types: morphologic descriptions (e.g., QT prolongation or widened QRS) and interpretations of those findings (e.g., myocardial infarction, atrial fibrillation, or ventricular hypertrophy). While many have developed automated feature extraction programs based on ECG waveforms, automated algorithms are imperfect, with accuracies of 42-96% [8, 9]. These algorithms are generally superior for morphological descriptions than their interpretations [8, 9]. However, many factors, such as an arrhythmia or ischemia, can alter the accuracy of morphological descriptions as well. For these reasons, cardiologists' interpretations of ECGs remain the consensus gold standard [8, 10, 11].

Natural language processing and concept-based indexing to standardized vocabularies such as the Unified Medical Language System (UMLS) have been applied to radiology reports [12, 13], clinical notes [14, 15], and medical education documents [16], among others. Previously, we reported the use of the KnowledgeMap concept indexer (KMCI) to identify UMLS concepts from cardiologist-generated ECG impressions [17]. KMCI is a general purpose concept identifier, using rigorous, scored-based algorithms to identify concepts from free text [16]. It accurately identifies unknown abbreviations, acronyms, and underspecified concepts (e.g., document phrase "1st degree block" for the closest UMLS match "1st degree *atrioventricular* block"). KMCI scores ambiguous concept matches using the context of other concepts matching around it to favor candidates that are likely to co-occur. Previously, we optimized and evaluated its performance on ECG impressions, finding a recall of 0.90 and precision of 0.94 [17]. It was especially accurate for myocardial perfusion changes, ECG rhythms, and extracardiac manifestations (recall and precision in excess of 0.98). This system, however, did not have the ability to detect negated or possible findings.

Table 1 – Comparison of KnowledgeMap Concept Identifier (KMCI) identification of negation to gold standard physician review

	Gold standard		
KCMI	Negated or possible findings	Positive or Probable findings	**Total concepts**
Identified as negated or possible	722 (6.9%)	13 (0.1%)	735 (7.0%)
Identified as positive	20 (0.2%)	9725 (92.8%)	9745 (93.0%)
Total Concepts	742 (7.1%)	9738 (92.9%)	10480

In this paper, we report on the integration of negation detection algorithms into KMCI and its initial application on a four-year collection of ECGs to identify QT prolongation. The ultimate goal is a codified database of ECG impressions for the development of a medication decision support system.

Methods

Creation of ECG database

Vanderbilt University Medical Center has developed an anonymized database of all orders, laboratory results, and ECGs for all inpatients admitted for 2-30 days from 1999-2003 as part of an ongoing research study investigating drug effects. The ECGs were imported in an XML format from an ECG management system. Every ECG includes machine calculated intervals and heart rate as well as a cardiologist-generated free-text impression. Cardiologists create an impression for all ECGs by selecting among personalizable stock phrases ("normal sinus rhythm") and editing as necessary ("normal sinus rhythm *with rare PVCs*"), or typing comments *de novo* ("LA abnormality, PVCs, and inferolat ST-T changes"). Finally, cardiologists code each ECG with a standard severity: normal, otherwise normal, borderline normal, or abnormal. We extracted all ECGs from our research repository and loaded them into a relational database. There were 44,808 ECGs in the database with more than 155,000 sentences.

From this dataset, we randomly selected a test set of about 5,000 sentences for development of our negation algorithm. Another 5,000 randomly selected sentences were reserved as a validation set. All subsequent data analysis was performed on the entire dataset of ECGs.

Negation tagging algorithm and evaluation

Many authors have developed negation tagging algorithms [18-20]. We applied a modified version of the NegEx algorithm [18] using regular expressions. We marked phrases as "negated", "possible", or "asserted." For the purposes of decision support, we included findings indicated as "probable" or "likely" as asserted findings. We used a total of 205 phrases indicating negation or possibility, including symbols such as "?" and "r/o" as indictors of "possible." We used a window of 8 words before or after a negating phrase. In this dataset, we found that most periods marked abbreviations rather than sentences. Semicolons, unmatched closing parentheses, and other negating phrases terminated the current negation window.

One author (JP, an internist), unfamiliar with the NegEx algorithm or its implementation, scored 5000 sentences from randomly selected ECGs via a color-coded HTML interface that highlighted the negating phrase and words modified by them. Only medical concepts or medical modifiers were considered for scoring. Concepts were marked as a correctly identified negated concept (i.e., a true positives, TP), false positives (FP), true negatives (TN), and false negatives (FN). We calculated recall of negated concepts as TP/(TP+FN); precision as TP/(TP+FP); and negative predictive value, the probability that the concept is not negated (i.e., an asserted finding), as TN/(TN+FN). Following the evaluation, three new negation phrases were added and validated over several thousand sentences with an improvement in matching before application to the entire dataset.

Development of concept-based ECG database

We applied KMCI to identify Unified Medical Language System (UMLS) concepts from the free text ECG impres-

Table 2 – Concepts employed to identify QT prolongation within ECG impressions.

Concept query (Asserted and Possible Concepts only)				**ECG matches**	**True Positives**
1	C0151878 Prolonged QT interval			2294	2193
2	C1560305 Prolonged QTc interval			55	55
3	C0023976 Long QT syndrome			2	2
4*	C0429028 QT interval C0489625 QTc interval C0860814 QTc	with any of	C0205166 Long C0439590 Prolonged C0392744 Lengthened	100	100
Total unique ECGs				**2,364**†	**2,363**†

*For query 4, one concept from each list must occur within the same sentence to be considered a match.
†Some ECGs matched more than one concept query (e.g., "QT prolongation. Compared with [date], the QT has lengthened.")

Table 3 – Comparison of QT prolongation identified in ECG impression to automated QTc by ECG machine.

	Concept query	QTc > 400	QTc > 450	QTc > 500	QTc > 550
ECGs matching criteria	2,364	34,059	11,804	2,518	620
KMCI positives	2,364	2,357	2,304	539	117
Sensitivity		1.00	0.98	0.23	0.05
Specificity		0.19	0.77	0.95	0.99
Positive predictive value		0.06	11,804	2,518	620
Negative predictive value		1.00	0.20	0.21	0.19

sions, using the optimizations identified in prior study [17]. We added synonyms and derivational transformations to KMCI's lexicon and modified the sentence-identification algorithm to ignore most spaces and periods when determining sentence breaks. We used the 2006AC version of the UMLS [21]; the only restriction on concept matching is favoring underspecified concepts with words such as "heart" and "electrocardiogram" (see [17] for a full list). Candidate UMLS concepts with these words are penalized less than candidates with other words when the words do not match a document word. We applied the negation algorithm following the concept identification step to mark each concept as positive, possible, or negated. The concept-identified ECGs are linked to the original ECG impressions and calculated intervals, forming the identified ECG dataset.

Analysis of QT prolongation

Through perusal of the ECG dataset, the authors identified the UMLS concepts representing "QT prolongation," including any text indicating a probable or possible QT or QTc prolongation. To verify that we had found all concepts representing QT prolongation, we also did text searches for all string matches with "QT" or "QTc." Each was manually verified by analyzing all unique strings containing the concept. We extracted these concepts, ignoring negated concepts, along with the automated QT and QTc intervals identified by the ECG management system. We compared the predictive value of the computer calculated QT and QTc intervals (a continuous number) via the area under the receiver operator characteristic curves (AUC) using the cardiologist interpretation as the gold standard and also by the positive predictive value of the commonly-used cutoff values of 400ms, 450ms, 500ms, and 550ms.

To assess the negative predictive value of our query, we reviewed a random sampling of 100 ECGs with calculated QT or QTc intervals > 450 for potential reasons why the QT calculations may be incorrect and to ascertain if any positive ECGs were missed by our concept query. Since the ECG images are not stored in our anonymized database, we were only able to evaluate the ECG intervals and raw text of the cardiologist impression. Student t-tests were used to compare parametric data. AUC and statistical analyses were performed with Stata, version 9.2 (StataCorp LP, College Station, TX).

Description of identified ECG database

To investigate the possibility for the database for decision support, we developed preliminary queries for a number of diagnoses that may be of interest for decision support (see Table 4). Each of these topics requires an interpretation of the free text. Topics were selected by finding the UMLS concepts representing the topic of interest in the database of matched concepts. For myocardial infarction, this involved the tree of concepts related to "myocardial infarction" and "infarct."

Results

Table 1 shows the results of the negation analysis. The 5,000 sentences in the negation test set contained a total of 10,480 UMLS concepts. Overall recall was 0.973 and overall precision 0.982. The negative predictive value of finding negation (probability that a statement was positive given it was identified as positive) was 0.998. All false negatives were due to three phrases not present in the regular expression list: "replaced <negated concept>", "<negated concept> replaced by", and "<negated concept> is/are gone." Several of the false positives were instances in which negating phrases were amid multiple concept words (e.g., "ST *no longer* depressed," in which the negated concept "ST depression" is separated by the negation phrase "no longer"). KMCI typically identified the correct UMLS concept for these phrases. Misspellings also caused some errors.

KMCI identified 375,838 concepts from the 44,080 ECGs in the database mapping to 23,080 unique admissions. Cardiologists identified 70% percent of the ECGs as "abnormal," 12% as "borderline," and 18% as "normal" or "otherwise normal." Of identified concepts, 339,554 (90.3%) were asserted, 29107 (7.7%) possible, and 7177 (1.9%) negated.

Table 2 shows the concepts used for the QT prolongation query and frequency of each in the database. There were 254 unique strings with a range of 2-18 words (median 11 words, weighted median 5 words) matching the QT prolongation query; 15 of these (e.g., "QT interval long for rate") accounted for nearly 90% of all matching impressions. The overall precision was 2363/2364 (1.00). Table 3 shows the results of different methods of predicting prolonged QT intervals; 2,364 ECGs (5.3% of all ECGs) were identified as representing QT prolongation by our concept query. The average QTc interval for those with prolonged QT intervals was 487 ms (range 363-716 ms); ECGs without mention of QT prolongation averaged 429 ms ($p<0.001$, range 46-785 ms). Overall, the calculated QT interval had an AUC of 0.73 for predicting QT prolonga-

Table 4 – Number of ECGs expressing potential targets for decision support

Concept	Total (%)	Positive	Possible	Negated
Myocardial infarction (MI)	6,355 (14)	3,381	2,919	55
Acute MI	208 (0.5)	167	36	5
Myocardial ischemia	2,312 (5.2)	1,855	444	13
ST segment elevation	1,015 (2.3)	895	92	28
Wolff-Parkinson-White bypass tract	107 (0.2)	93	7	7
Atrioventricular block, 1st degree	2,461 (5.5)	1,876	571	14
Atrioventricular block, 2nd degree	61 (0.1)	58	1	2
Atrioventricular block, 3rd degree	24 (0.05)	23	1	0
Pericarditis	105 (0.2)	43	62	0
Atrial fibrillation	1,748 (4.0)	1,719	16	13
Atrial flutter	397 (0.9)	373	20	4
Total number of ECGs	**44,080**			

tion; the QTc interval AUC was 0.91. Of the 100 manually-reviewed ECGs with QTc intervals longer than 450 ms but that were negative by concept query, 32% had a bundle branch block; 24% had various ST segment or T wave abnormalities; 24% had an arrhythmia, aberrant complexes, or a pacemaker; and 23% had myocardial ischemia or infarct. No ECGs contained comments suggesting QT prolongation. Only 4 ECGs had no significant electrocardiographic abnormalities that could not alter calculation of the QT interval.

Table 4 shows concept findings over the entire database.

Discussion

We studied the application of a concept-based, natural language processing system to identify QT prolongation within cardiologist-generated ECG reports. Commonly used cutoffs for diagnosing QT prolongation from the QTc intervals calculated by ECG machines were poor, with positive predictive values of only 6-21% when compared with the NLP-based approach. Physician review of ECGs with long QTc by calculation but not by the NLP query found 94% of these ECGs had findings that would confound automatic QTc calculation. No ECGs not identified by KMCI were identified as prolonged by the cardiologist. KMCI is a more accurate means of identifying QT prolongation than automated interval analysis. Due to the high provider override rates in most medication decision support systems, due in part to poor specificity [22], a medication decision support system for QT prolongation requires use of the cardiologist-generated impression.

A modified version of the NegEx negation algorithm performed well in detected negation within this dataset with a recall of 97.3% and a precision of 98.2%. Negation within ECGs was rare; overall only 1.9% of all concepts were marked as negated and 7.7% as possible. Thus, the probability of a concept the system identified as positive truly being positive was 0.998. The high recall and precision of negation detection in this dataset is likely due to a constrained vocabulary and the relative simplicity of the ECG impression sentences compared to prior studies in other clinical document types, which had recalls of 78-97% and precisions of 85-91%[18-20].

We have grouped terms such as “questionable”, “cannot rule out”, and “borderline” together. A more granular approach to negating terms may be more predictive of specific clinical outcomes. Also, negating phrases such as “no longer” indicate that a patient has a history of a finding as well as the absence of it now; the current algorithm only identifies the latter. Such information may help determine treatment efficacy. Finally, we used a rather simple “negation window” technique consisting of a certain number of words before or after a negation phrase. Some commas separate clauses while others represent a list of concepts. A more advanced algorithm could use a parsed sentence and the presence of prepositions or coordinating conjunctions to correctly size of the negation window.

The handling of “possible” findings, included in our QT prolongation query, would vary depending on application. For the purposes of decision support, inclusion of potential findings may help prevent adverse events. For example, one would want to avoid starting a drug known to prolong the QT interval in a patient that had a “borderline long QT.” In addition, many uncertain ECG findings require further workup. A patient with potential ischemia requires further evaluation, and one would likely discontinue cyclooxegenase-2 inhibitors in this patient. For research investigations, however, one may desire to exclude uncertain findings, as many ECGs indicating “cannot rule out” may represent benign changes.

The next step in this investigation involves building a decision support tool. Many resources list possible QT prolonging medications, such as the medication registry maintained at the University of Arizona [7]. A decision support system for QT prolongation could intercept orders for high-risk medications. By investigating the occurrences of medication orders in our retrospective database of inpatient admissions in individuals with QT prolongation, we may be able to identify potential medications interactions and new causes of QT prolongation. Some of these interactions may be indirect, such as the addition of a

potent cytochrome P450 inhibitor that raises serum concentrations of a known offending agent. Ideally, a medication intervention could not only intercept medications that prolong the QT interval but also those that significantly interact with those already prescribed.

Since this is a full-text concept index over a large dataset of ECGs, we also have the ability to support many other types of decision support and research venues. While a string-matching algorithm could potentially replicate the performance of the QT prolongation concept query, it would lack flexibility and scalability. By fully concept indexing, we can quickly assess multiple queries, enabling a broader range of research and decision support tools. In addition, the parent-child relationships between concepts in the UMLS ameliorate querying across broad concepts, such as myocardial infarction.

In this study, we used a general purpose concept-identification program with the entire UMLS. We optimized the algorithm to enhance synonymy and favor underspecified matches that match cardiology-related concepts. By processing these ECGs in bulk, KMCI is able to "learn" a corpus of frequently-occurring concepts and thus favors those concepts and their related concepts when encountering ambiguous matches. Other general purpose concept identification algorithms may not perform as well without similar optimizations.

The interpretation of our findings is limited. The performance of the negation algorithm and concept identifier may not translate to other repositories of ECG impressions; however, no specific optimizations were made that correspond to our institution. Second, while we evaluated the QT prolongation algorithm against a random set of ECGs for false negatives, we did not have access to the original ECG waveforms to reevaluate for potential false negatives due to errors in the cardiologists' reading; however, we expect that this would be unlikely to dramatically alter our results. Furthermore, the prevalence of QT prolongation in the KMCI-identified dataset (5.3%) carries more face validity than the QTc interval prevalence of 26.6%. Third, since cardiologists read all ECGs in our institution in a timely manner, decision support based on the textual interpretation of ECGs is feasible. It may not be practical at other institutions in which formal ECG impressions are not rapidly available. Fourth, while we have accurately identified concept matches and their negation status, this is not the same as asserting normality. Our algorithm tells the presence or absence of "atrial fibrillation," for instance, but cannot tell that there were no arrhythmias. These questions may be addressed by classifying concepts by type (e.g., "rhythm" or "perfusion abnormalities") and defining normal status. Finally, our exploratory list of concepts in Table 4 has not been formally assessed for false negatives and provides only a rough incidence of these findings in set of ECGs.

Conclusion

The combination of a negation tagging algorithm with an effective concept identifier allows rapid assessment of clinical syndromes such as QT prolongation, myocardial ischemia, or atrioventricular conduction disturbances. We believe this technique may enable large-scale research on drug adverse events and development of new decision support tools to improve cardiovascular medication safety.

Acknowledgements

We would like to thank Mr. Mark Arrieta for his assistance in assembling the database of ECGs. This work is supported by two National Library of Medicine grants, T15 LM007450 and R01 LM007995.

References

[1] Ray WA, Murray KT, Meredith S, Narasimhulu SS, Hall K, Stein CM. Oral erythromycin and the risk of sudden death from cardiac causes. N Engl J Med 2004 Sep 9; 351(11): 1089-96.

[2] Roden DM. Drug-induced prolongation of the QT interval. N Engl J Med 2004 Mar 4;350(10):1013-22.

[3] Bresalier RS, Sandler RS, Quan H, Bolognese JA, Oxenius B, Horgan K, Lines C, Riddell R, Morton D, Lanas A, Konstam MA, Baron JA. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med 2005 Mar 17;352(11):1092-102.

[4] Solomon SD, McMurray JJ, Pfeffer MA, Wittes J, Fowler R, Finn P, Anderson WF, Zauber A, Hawk E, Bertagnolli M. Cardiovascular risk associated with celecoxib in a clinical trial for colorectal adenoma prevention. N Engl J Med 2005 Mar 17;352(11):1071-80.

[5] Zeltser D, Justo D, Halkin A, Rosso R, Ish-Shalom M, Hochenberg M, Viskin S. Drug-induced atrioventricular block: prognosis after discontinuation of the culprit drug. J Am Coll Cardiol 2004 Jul 7;44(1):105-8.

[6] Bazett H. An analysis of the time relationship of electrocardiograms. Heart. 1920;7:353-70.

[7] Drugs that prolong the QT interval and/or induce Torsades dr Pointes ventricular arrhythmia. [cited 2006 12/2]; Available from: http://www.arizonacert.org/medical-pros/drug-lists/drug-lists.htm

[8] Tsai TL, Fridsma DB, Gatti G. Computer decision support as a source of interpretation error: the case of electrocardiograms. J Am Med Inform Assoc. 2003 Sep-Oct; 10(5):478-83.

[9] Willems JL, Abreu-Lima C, Arnaud P, van Bemmel JH, Brohet C, Degani R, Denis B, Gehring J, Graham I, van Herpen G, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. N Engl J Med 1991 Dec 19;325(25):1767-73.

[10] Olsson SE, Ohlsson M, Ohlin H, Dzaferagic S, Nilsson ML, Sandkull P, Edenbrandt L. Decision support for the initial triage of patients with acute coronary syndromes. Clin Physiol Funct Imaging 2006 May;26(3):151-6.

[11] Paoletti M, Marchesi C. Discovering dangerous patterns in long-term ambulatory ECG recordings using a fast QRS detection algorithm and explorative data analysis. Comput Methods Programs Biomed 2006 Apr;82(1):20-30.

[12] Fiszman M, Chapman WW, Aronsky D, Evans RS, Haug PJ. Automatic detection of acute bacterial pneumonia from chest X-ray reports. J Am Med Inform Assoc 2000 Nov-Dec;7(6):593-604.

[13] Huang Y, Lowe HJ, Hersh WR. A pilot study of contextual UMLS indexing to improve the precision of concept-based representation in XML-structured clinical radiology reports. J Am Med Inform Assoc 2003 Nov-Dec;10(6):580-7.

[14] Friedman C, Shagina L, Lussier Y, Hripcsak G. Automated encoding of clinical documents based on natural language processing. J Am Med Inform Assoc 2004 Sep-Oct;11(5):392-402.

[15] Nadkarni P, Chen R, Brandt C. UMLS concept indexing for production databases: a feasibility study. J Am Med Inform Assoc 2001 Jan-Feb;8(1):80-91.

[16] Denny JC, Smithers JD, Miller RA, Spickard A, 3rd. "Understanding" medical school curriculum content using KnowledgeMap. J Am Med Inform Assoc 2003 Jul-Aug;10(4):351-62.
[17] Denny JC, Spickard A, Miller RA, Schildcrout J, Darbar D, Rosenbloom ST, Peterson JF. Identifying UMLS concepts from ECG Impressions using KnowledgeMap. AMIA Annu Symp Proc 2005:196-200.
[18] Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform 2001 Oct;34(5):301-10.
[19] Elkin PL, Brown SH, Bauer BA, Husser CS, Carruth W, Bergstrom LR, Wahner-Roedler DL. A controlled trial of automated classification of negation from clinical notes. BMC Med Inform Decis Mak 2005;5(1):13.
[20] Mutalik PG, Deshpande A, Nadkarni PM. Use of general-purpose negation detection to augment concept indexing of medical documents: a quantitative study using the UMLS. J Am Med Inform Assoc 2001 Nov-Dec;8(6):598-609.
[21] Unified Medical Language System. [cited 2006 12/2]; Available from: http://www.nlm.nih.gov/research/umls/
[22] van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc.2006 Mar-Apr;13(2):138-47.

Address for correspondence:

Josh Denny, MD
4th floor, Eskind Biomedical Library
2209 Garland Ave
Nashville, TN 37232, USA
josh.denny@vanderbilt.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

A Comparison of Supervised Classification Methods for Auditory Brainstem Response Determination

Paul McCullagh, Haiying Wang, Huiru Zheng, Gaye Lightbody and Gerry McAllister

Department of Computing and Mathematics, University of Ulster, United Kingdom

Abstract

The ABR is commonly used in the Audiology clinic to determine and quantify hearing loss. Its interpretation is subjective, dependent upon the expertise and experience of the clinical scientist. In this study we investigated the role of machine learning for pattern classification in this domain. We extracted features from the ABRs of 85 test subjects (550 waveforms) and compared four complimentary supervised classification methods: Naïve Bayes, Support Vector Machine, Multi-Layer Perceptron and KStar. The ABR dataset comprised both high level and near threshold recordings, labeled as 'response' or 'no response' by the human expert. Features were extracted from single averaged recordings to make the classification process straightforward. A best classification accuracy of 83.4% was obtained using Naïve Bayes and five relevant features extracted from time and wavelet domains. Naïve Bayes also achieved the highest specificity (86.3%). The highest sensitivity (93.1%) was obtained with Support Vector Machine-based classification models. In terms of the overall classification accuracy, four classifiers have shown the consistent, relatively high performance, indicating the relevance of selected features and the feasibility of using machine learning and statistical classification models in the analysis of ABR.

Keywords:

Auditory Brainstem Response, wavelet decomposition, feature extraction, classification, decision support

Introduction

The Auditory Brainstem Response (ABR) is evoked when a stimulus click is applied to a subject's ear to determine hearing acuity and integrity of the auditory pathways. If the stimulus is perceived, a response changes their electroencephalogram (EEG) within 10ms from stimulus onset (Figure 1). The amplitude of the ABR signal is approximately 1 µVolt -5µVolts and is hidden behind the background EEG and noise (approximately 50µVolts). The components of the ABR are swamped by the endogenous electrical activity of the brain and the determination of a response can be difficult particularly at low levels of acoustic stimulation, as hearing threshold is reached. The ABR waveform is extracted by coherent averaging which exploits the deterministic nature of the signal to enhance the waveform while suppressing the uncorrelated EEG, extraneous noise and artifact. It is necessary to average approximately 1000-2000 trials before the noise is sufficiently suppressed, with signal to noise ratio enhanced proportional to the square root of the number of trials.

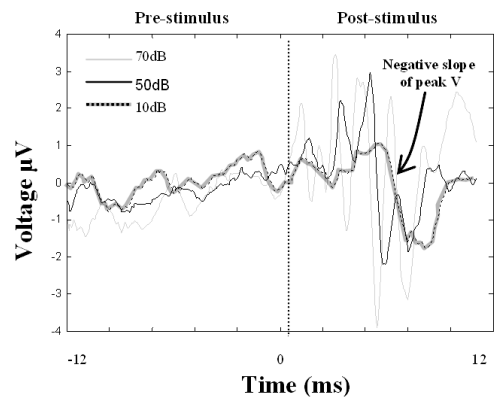


Figure 1 - ABR waveform at a 70, 50 and 10 dB Stimuli

Figure 1 shows responses for a healthy adult ranging from 70dB (normal hearing level, nHL) stimulus to threshold (10dB nHL). The characteristic shape is referred to as the Jewett waveform, comprising up to seven peaks, labeled I-VII. As the level of stimulus is reduced, the different peaks of the waveform become less defined, and their latency is increased. The shape of the waveform from the long slope at the top of peak V is the strongest part of the waveform to remain as the stimulus diminishes. When the stimulus level is set near the subject's hearing threshold the identification of wave V and its following negative slope assist with classifying the presence of the response. The shape of the ABR response will differ between subjects depending on a number of factors such as, electrode placement, filter settings, stimulus intensity, ear used, age, sex and subject's skull thickness. For this reason the range of factors need to be taken into consideration before a clinical expert can make an interpretation. In particular the expert checks the latencies of waves I, III and V and examines the morphology of the waveform. Often sub-averages based on successive recording sessions are compared for consistency and a cross-correlation may also be taken between repeat recordings to assist in classification.

The interpretation of the ABR is subjective, thus clinical experts may not always draw the same conclusion [1], particularly at auditory threshold. Artificial intelligence methods have been used to provide objective assistance in response interpretation [2]. Useful information may be extracted from the EEG recordings using features extracted from the time and frequency domains [3]. Davey [4] showed that the ratio of the power of the post over pre stimulus time domain waveforms could be used as an effective method to classify 'strong' responses with an accuracy of 98.6%. Remaining ABR waveforms, i.e. those which did not have a strong response, were passed to a second stage of analysis, whereby repeat recordings were used to derive features based on cross-correlation parameters using both the time and frequency domains. Lower accuracies ranging from 65% to 82% were achieved, dependent upon features used. The lower accuracies reflect the more difficult classification process.

Lightbody [5] used wavelet decomposition [6, 7] on the same dataset following a two-stage classification process. Power ratios of the post over pre stimulus wavelet coefficients were used to classify 'strong' responses. Those remaining were then classified using correlation features of repeated tests derived from the wavelet domain. Overall accuracy of 76.4% was obtained using a C5.0 decision tree classifier. Strong responses were classified without error. The criterion to determine a 'strong' response required a combination of time and wavelet post stimulus to pre stimulus power ratios (time domain power>2 and wavelet domain power>1.6, determined heuristically). By combining features using the Demster-Shafer method, as used in evidential reasoning [8], Davey achieved a classification accuracy of 95.6% for 'strong' responses and 85% for lower level responses [9].

In this study we compared four additional classification techniques which are in widespread use by machine learning researchers. The aim was to determine whether one of these techniques provides superior classification accuracy for the quasi-stationary ABR evoked response. Three additional statistical descriptors, namely precision, sensitivity and specificity were used to further explore the data set. The data set comprised both high and low level recordings. A secondary aim was to determine whether a single ABR recording, irrespective of level, provided sufficient information for decision support. Hence the classification process does not involve correlation parameters from repeated recordings.

Methods

The study was performed on a database of 85 test subjects, provided by the Audiology Department of the Royal Group of Hospitals in Northern Ireland. Each subject had a range of test stimulus levels applied providing a mix of good, weak and non-response waveforms, all of which were classified by a clinical expert. There were 550 waveforms in total, 396 recordings with a YES classification and 154 with a NO classification.

The data were pre-processed by band-pass filtering (100Hz-3kHz), sampled at 20kHz and then de-noised using a wavelet filter. Between 1000 and 2000 ensembles were averaged to provide one ABR waveform. Each waveform consisted of 480 data samples, half before the stimulus and half after, which related to 12ms before the stimulus and 12ms after the stimulus. The ABR waveform appears within 10ms of stimulus onset, which after band-pass filtering and sampling relates to 200 data points. A number of features were extracted from both time and wavelet domains. The Daubechies wavelet has been used to de-noise biosignals [10].

We computed a range of scaling lengths performed to 6 levels (A6: 0–156Hz, D6: 156–312Hz, D5: 312–625Hz, D4:625-1250Hz, D3:1250-2500Hz, D2:2500-5000Hz, D1:5kHz-10kHz). Levels D6, D5 and D4 showed to be of significant interest as they related to the key frequencies (200, 500 and 900Hz [11]) contained within the ABR. This is depicted in figure 2 which shows the Fast Fourier Transform (FFT) amplitude values for post-stimulus ABR waveforms at 70dB and 30dB stimulation intensities. It highlights the frequencies covered by the D6, D5 and D4 wavelet coefficients.

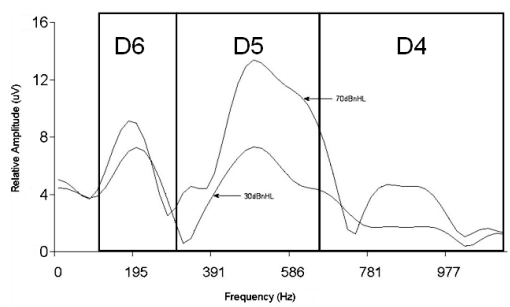


Figure 2 - FFT amplitude values for post-stimulus ABR waveforms at 70dB and 30dB stimulation intensities

Although in figure 2 the D4 coefficients seemed to show the least promising power peaks they have double the number of coefficients than the D5 range helping to reveal more feature consistency. Figure 3 illustrates the D4 component for pre and post stimulus activity for decomposition on 256 data points.

The decomposition was carried out on the full pre stimulus section and only waves I to V of the post section. Davey [4] showed that by using data from 1.5 to 9.5 ms post stimulus then the classification accuracy was improved. By focusing in on the ABR where waves I to V where most likely located removed the samples early in the waveform that may be more likely to contain artifacts and influence the classification. This same process was used for the wavelet decomposition of the post stimulus section. For both the pre and post stimulus waveforms data extension was required to extend the dataset to 256 which is the nearest dyadic number [5] so to support wavelet decomposition. Different methods to extend the data had been looked at to determine the method least likely to incur boundary issues [12].

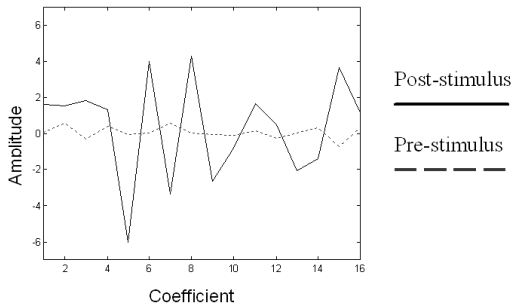


Figure 3: Extracted D4 wavelet coefficients (70dB stimulus), for pre and post-stimulus regions

An alternative partitioning of the data has been considered by performing wavelet decomposition on a smaller section of the data containing 128 data points, therefore removing the need for data extension [13]. The window of data relates to a 6.4ms section of the waveform, so the choice of window within the full post stimulus section is vitally important to ensure that the subject's response has been captured. This is crucial because as the strength of a response diminishes the peaks fade and latencies alter until peak V remains. Also, the exact position of the response differs from subject to subject. A sliding window of 128 data points was used to derive features between 1.5ms and 9.5ms of the post stimulus waveform, which was shown to be the time frame with the strongest waveforms in an ABR response [4].

Using a moving window of 128 samples from stimulus onset, the window containing the peak time domain power ratio was determined. The ABR response is expected to be at its maximum within 1.5ms and 9.5ms post stimulus. It was considered that a peak power before the 1.5ms mark could signify that the power ratio might be a result of something other than the ABR. A weighting based on the position of the peak power was used to numerically represent this doubt in the power ratio as a feature for classification. This weighting was applied to features DataAbs and DataPwr resulting in features wDataAbs and wDataPwr respectively. Whereby, DataAbs is a feature derived from the ratio of the post and pre stimulus D4 coefficients [5] (Figure 3) and DataPwr is the ratio of the post and pre stimulus time domain waveform [4]. In both cases only waves I to V in the post stimulus waveform were considered and the wavelet decomposition was performed on 256 data points resulting in 16 D4 coefficients (Figure 3).

Each 128 sample window starting from sample 35 to starting from sample 75 were independently analyzed. Power ratios were calculated labeled RNxx, where R represents the 'response' post stimulus region in the waveform and N represents the 'no response' pre stimulus region of the waveform, and 'xx' represents the starting sample for the window of data. Similar ratios for the D4, D5 and D6 wavelet coefficients were calculated for each of the 128 sample windows. Using all the D4-D6 coefficients was first considered since a 6-level decomposition of a 128 sample waveform results in a small number of coefficients (8-D4, 4-D3 and 2-D6 coefficients). Then the individual bands were considered. The last features were derived by averaging the calculated features for each of the 128 sample windows, so to capture any overall trends.

The features were used as inputs to four classification models Naïve Bayes (NB), Support Vector Machine (SVM), Multi-Layer Perceptron (MLP) and KStar. NB is a probalilistic classifier based on the Bayes rule of conditional probability. It assumes independence between features. It uses the normal distribution to model numeric attributes by calculating the mean standard deviation for each class [14]. SVM is a kernel based classifier. The basic training for SVMs involves finding a function which optimizes a bound on the generalization capability, i.e., performance on unseen data. By using the kernel trick technique, SVM can apply linear classification techniques to non-linear classification problems [15]. A MLP is a non-linear classification approach that may be trained using the back propagation algorithm. A MLP consists of multiple layers of computational units (an input layer, one or more hidden layer and one output layer). For the MLP model, the results were obtained using a model consisting of one hidden layer with six nodes when evaluating the top ten features, four nodes when evaluating the top six features and two nodes when considering the top three features (the choice of feature subsets is discussed later in the paper). Each MLP was trained for 500 epochs and the learning rate was set to 0.3. The final classification is KStar, a relatively simple instance based method. The class of a test instance is based upon the class of those training instances similar to it, using an entropy-based distance function to compute the similarity between two cases. The underlying assumption is that similar instances will have similar classes. All four classification models were implemented within the framework provided by the open-source platform Weka package [16]. This provides a collection of machine learning algorithms for data mining tasks, which may be implemented directly or called from Java code.

In the evaluation of each classifier we used ten-fold cross validation, i.e. the entire dataset is partitioned into 10 subsets of approximately equal size. Each subset in turn is used as the test set while the other subsets are combined to form a training set. The quality of each classifier was assessed by the extent to which the correct class labels have been assigned. Each classifier is assessed by three statistical measures: Precision (Pr), Sensitivity (Se), Specificity (Sp). All the features were ranked using correlation coefficient-based ranking.

Results

The top five features are illustrated in Table 1. Table 2 details accuracy, precision, sensitivity and specificity of the classifiers in separating the two classes, based on the top 5 discriminative features. (Note as this is a two-class problem, Se and Sp are reversed for Response='NO' and Response='YES').

Table 1 - Top five predictive features

Feature	Name	Description
1	wDataAbs	Wavelet power feature: D4 (weighted)
2	DataAbs	Wavelet power feature: D4
3	RN75	Power ratio: 128 samples from 75
4	RN70	Power ratio: 128 samples from 70
5	RN65	Power ratio: 128 samples from 65

Table 3 details accuracy, precision, sensitivity and specificity of the classifiers in separating the two classes, based on the top 2 discriminative features. The accuracy and precision in determining a NO response are slightly reduced, implying that most discrimination is harnessed from the relative power measure from the wavelet domain, i.e. the D4 coefficient. Time may have a small role to play in the overall morphology.

Table 2 - Prediction results for four classifiers using 10-fold cross validation with top 5 features (wDataAbs, DataAbs, RN75, RN70, and RN65)

Method	Ac (%)	Response='NO'			Response='YES'		
		Pr (%)	Se (%)	Sp (%)	Pr (%)	Se (%)	Sp (%)
NaiveBayes	83.4	93.5	82.2	86.3	66.8	86.3	82.2
SVM	81.9	83.2	93.1	54.9	76.8	54.9	93.1
MLP	81.9	87.9	86.2	71.4	68.3	71.4	86.2
Kstar	82.2	86.9	88.1	68.0	70.4	68.0	88.1

Table 3 - Prediction results for four classifiers using 10-fold cross validation with top 2 features (wDataAbs and DataAbs)

Method	Ac (%)	Response='NO'			Response='YES'		
		Pr (%)	Se (%)	Sp (%)	Pr (%)	Se (%)	Sp (%)
NaiveBayes	81.4	86.7	86.9	68.0	68.4	68.0	86.9
SVM	80.2	81.5	93.1	49.1	74.8	49.1	93.1
MLP	80.7	84.5	89.1	60.6	69.7	60.6	89.1
KStar	81.2	85.0	89.1	62.3	70.3	62.3	89.1

Discussion

The ABR test is routinely performed to detect hearing loss (Response='NO') in the Audiology clinic. Where the ratio of post-stimulus power to pre-stimulus power is high (>5), the classification decision can be made to high accuracy (98.6%). This is manifest as a clear Jewett response. Misclassification may be due to artifact attributed to stimulus, myogenic (muscle) activity, or eye blink artifact which corrupts the ABR. Such activity may be detected by an expert due to latency; stimulus artifact occurs within the first msec post-stimulus, and myogenic artifact normally occurs after wave VII. Eye blinks may be harder to eliminate as they may be more randomly distributed in the pre and post stimulus activity and hence overlap with the ABR components. The strategy is normally to remove eye blinks at source, as much as possible, by rejecting a contaminated trial.

ABR responses with a lower post-stimulus power to pre-stimulus power ratio (<5), provide a more difficult classification task, with a classification accuracy of 76.4% reported using the C5.0 decision tree classifier. This second stage classification process required the use of correlation features, from repeated trials, to assess consistency of features. In the context of decision support this would provides a complex implementation. In this research we have utilized the same data set, to assess four additional classifiers, used by data mining researchers. The classification is based only on single ABR waveforms, which provides a more difficult classification task, but is more straightforward to implement. In the context of decision support, a classifier with high accuracy and high sensitivity is required. The best overall accuracy of 83.4% was obtained for the statistically based Naïve Bayes method, using the top five features. This has a good balance between sensitivity (82.2%) and specificity (86.3%). The SVM classifier has a higher sensitivity (93.1%), but at the expense of a lower specificity (54.9%). The SVM also maintains its sensitivity as the number of features reduces to two.

In terms of mathematical complexity, KStar, an instance-based classifier, and Naive Bayes (NB), a Bayesian theorem-based model, are relatively simple classifiers, but KStar requires additional space to store training datasets. In the case where the dataset is large, using KStar could be computation-intensive. In terms of algorithm configuration, the implementation of NB is quite straightforward. It does not require any predefined parameters. MLP and SVM have more complicated parameters settings to tune. For example, MLP requires users to specify the learning rate and the number of learning epochs in advance; and SVM requires users to predefine an appropriate kernel function. NB has been proven to be one of most proficient classification algorithms in most cases. Nevertheless, given the size of datasets studied and current advanced computational resources, the differences between four classifiers in terms of computational time are not obvious in this study. Each classifier has its own advantages. For example, NB can achieve the best results in terms of specificity (86.3%). The best sensitivity was obtained by using SVM-based classification models (93.1%). Therefore, it is

hard to say which one is best in this case. Nevertheless, it is expected that a better result could be obtained by combining different classifiers. A subsidiary aim of this study was to assess the suitability of the classifiers in evoked potential analysis. However there was little difference in the overall accuracy and the choice of classifier may be more influenced by ease of implementation. A further analysis is required to assess the agreement between the classifiers, and this can be undertaken by analyzing the classification/ misclassification of single response. This study illustrates the need for further validation of the approach by applying unseen data set to the classifiers.

Conclusions

The overall classification accuracy is quite similar for the four methods, ranging from 81.9% to 83.4% with the top 5 features and 80.2% to 81.4% with the top 2 features. These accuracies exceed the C5.0 (76.4%) determined in a previous study, but may be inflated because 'strong' responses are also included in this study as we adopt classification on a single averaged trail. A valid between-study comparison would require a 2 stage classification, as used in the previous study. Currently learning parameters required for each classifier are based on a trial-and-error approach. Combining other machine learning techniques such as the genetic algorithm to automatically determine the optimal learning parameters would enhance our methodology. All the classifiers show the consistent high accuracy which indicates the relevance of selected features and the feasibility of using machine learning and statistical classification models in the analysis of ABR. Additionally the feature ranking technique used is not optimal and could be further enhanced.

Acknowledgment

The authors wish to thank Mr HG Houston, Royal Group of Hospitals for providing classification on the ABR data set.

References

[1] Pratt TL, Olsen WO. Four-channel ABR recordings: Consistency in interpretation. Am J of Audiology 1995, 4 (2): 47-54.

[2] Delgado RE and Ozdamar O. Automated Auditory Brainstem Response Interpretation. IEEE Engineering in Medicine and Biology 1994, 227-237.

[3] Aoyagi M, and Harada J. Application of fast Fourier transform to auditory evoked brainstem response. Med. Inform., 1988, 13(3), 211-220.

[4] Davey R, McCullagh PJ, McAllister HG, Houston HG. The use of artificial neural networks for the objective determination of hearing threshold, using the auditory brainstem response. In: Neural Networks in Healthcare: Potentials and Challenges, IDEA Group Inc, 2006.

[5] Lightbody G, McCullagh PJ, and McAllister HG. Feature Extraction in Auditory Brainstem Responses: Use of the Daubechies Wavelet. EMBEC Conf 2005.

[6] Bradley A and Wilson W. On Wavelet Analysis of Auditory Evoked Potentials. J Clin Neurophysiology 2004, 115: 1114-1128.

[7] Mallat SG. A Theory for Multiresolution Signal Decomposition: the Wavelet Representation. IEEE Trans. on Pattern Anal. Machine Intelligence 1989, 11: 674-693.

[8] Shafer G. and Logan R. Implementing Dempster's rule for hierarchical evidence. Artificial Intelligence 1987, 33:271-298.

[9] Davey R, McCullagh P, Lightbody G and McAllister G. Auditory Brainstem Classification: a Hybrid Model using Time and Frequency Features, Artificial Intelligence in Medicine 2006. [Epub ahead of print].

[10] Zhang R, McAllister G, Scotney B, McClean S, Houston G. Combining wavelet analysis and Bayesian networks for the classification of auditory brainstem response. IEEE Trans Inf Tech Biomed 2006, Jul; 10(3):458-67.

[11] Ozdamar O, Delgado RE. Automated electrophysiologic hearing testing using a threshold-seeking algorithm. J Am Acad Audiol 1994, 5, 77-88.

[12] Gutierrez A, Somolinos A. Influence of wavelet boundary conditions on the classification of biological signals. Proceedings of the IEEE 26th Annual Northeast Bioengineering Conf 2000, 25-26.

[13] Lightbody G, McCullagh PJ, and McAllister HG. Feature extraction in auditory brainstem responses using wavelet decomposition on a moving window of waveform Data, IEE Irish Signals and System Conf 2006, 321-326.

[14] Irina R. An empirical study of the naive Bayes classifier. IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, 2001.

[15] Burges C. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery 1998: 2: 121-167.

[16] Witten IH and Frank E. Data Mining: Practical machine learning tools and techniques, 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

Address for correspondence

Dr PJ McCullagh,
School of Computing and Mathematics,
University of Ulster,
Shore Road,
Newtownabbey,
BT370QB,
N Ireland,
United Kingdom

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Non-linear Analysis for the Sleepy Drivers Problem

Ioanna Chouvarda[a], Christos Papadelis[a], Chrysoula Kourtidou-Papadeli[b], Panagiotis D Bamidis[a], Dimitris Koufogiannis[a], Evaggelos Bekiaris[c] and Nikos Maglaveras[a]

[a] *Lab of Medical Informatics, The Medical School, Aristotle University of Thessaloniki, Greece*
[b] *Greek Aerospace Medical Association, Thessaloniki, Greece*
[c] *Institute of Transport, CERTH, Thessaloniki, Greece*

Abstract

The problem addressed in this work is sleepiness during driving, which often leads to accidents in the streets. Experiments with sleepy drivers took place and the EEG data were analysed in terms of non-linear methods. Sample entropy and phase synchronization variations were investigated within the signal sections corresponding to "driving events", i.e. driving mistakes or loss of control, as well as to periods of drowsiness and sleepiness, as compared to the periods of normal driving. Decreased sample entropy, indicating loss of complexity, and an increased phase synchronisation have been found in the preliminary study presented. The results are encouraging towards developing an alerting system for predicting and preventing driving accidents.

Keywords:

sleep disorders, nonlinear dynamics, automobile driving

Introduction

Sleep is necessary for human beings and its gradual loss can lead from extreme short-term sleepiness to chronic sleepiness. Yet, the conditions of modern life increase stress and often lead people into sacrificing sleep despite the negative effects on health and abilities. In addition, as much as 20% of the general population suffers to some degree from sleep disorders which may cause extreme tiredness, loss of concentration and a pronounced inability to function normally, raising the risk of causing traffic and work accidents.

Sleepiness reduces reaction time, vigilance, alertness and concentration. Lack of adequate sleeps is associated with excessive fatigue, hypovigilance, stress and impairment of attention, information processing speed and decision-making quality, which are among the key causes of serious industrial accidents, including most nuclear accidents (Chernobyl, Three-mile Island) and other large scale accidents (such as the Bhopal chemical disaster) [1]. In addition, about 40% of fatal accidents on US highways are fatigue-related, whereas sleep-related accidents account annually for as many as 240,000 motor-vehicle accidents. UK statistics show that 25% of motor accidents are associated with driver fatigue [2]. According to NASA's Aviation Safety Reporting System, approximately 21% of the aviation incidents are fatigue-related.

As sensor, communication and artificial intelligence technologies are evolving, an effort is being paid to apply and combine such technologies in ambient intelligence systems for a more secure environment that detects and copes with the effects of sleepiness. An important challenge in this effort is the development of methods to define sleep/wakefulness involuntary transitions by physiological EEG measurements. Such predictions made feasible pervasively and unobtrusively can help decrease drastically the number of fatigue or sleepiness-related accidents and make work and streets a safer place. The European IST project SENSATION is involved in this field. A series of experiments concerning involuntary sleepiness have taken place, along with development of new miniaturized relevant sensors and new tools for processing the derived data in a meaningful way.

Within this scope, a sleepy driver experiment has taken place aiming at deriving new information regarding the driver's physiological status when losing driving control. In a previous work [3], frequency analysis of these experimental data lead to findings such as an increase of slowing activity and an acute increase of the alpha waves 5-10 seconds before driving events, accompanied by A rapid decrease of both Shannon and K-L entropies have been observed just before the driving events, i.e. loss of driving control.

A more extensive quantitative analysis is presented in this work, focusing on non-linear analysis. The hypothesis tested here that EEG "systemic" features, like complexity and synchronization, as calculated during incidences of sleepiness, and mainly during a driving event, decline from the ones calculated during driving without events. The detection of "patterns of sleepiness" among such features may help predict dangerous driving events.

Materials and methods

Subjects and experimental protocol

Five subjects participated in the present study. The subjects were average drivers (mean driving experience: 8.3 years), with a mean age of 26.5 years. The experiments

were performed in CERTH, Thessaloniki, Greece, from 6 June till 27 July 2005.

The subjects were asked to stay awake for at least 24 hours, and then to arrive at CERTH premises at around 22.00. Upon arrival and after passing the standard medical examination test, the subjects' level of sleepiness was estimated by using the Karolinska Sleepiness Scale (KSS) test, and their sleepiness behavior was scaled by the M.D. by using the Epworth Sleepiness test (1 = 'extremely alert', 5 = 'neither alert nor sleepy', 9 = 'extremely sleepy – fighting sleep'), resulting in two subjects with score 7 (very sleepy), two with grade 6, and one with 5 (neither sleepy nor awake).

The measurements were performed in the CERTH experiment car. The subject had to seat at the driver seat and the attached electrodes were connected to an ambulatory EEG monitoring system. EEG channels Fp1, C3, P3, O1, Fp2, C4, P4 and O2 were recorded, along with two horizontal and two vertical EOG channels for bipolar signals, two EMG channels and two ECG channels. For the data acquisition, a sampling rate of 200Hz, for each channel of the recording. The monitoring system hardware filters were adjusted to the band pass filtering option with a frequency range of 0.5 to 70Hz for EEG, with a notch filter at the 50Hz power supply component. A battery box with power supply independence of approximately 3 hours supported the monitoring system.

An experienced driving instructor was seated at the co-driver's seat, as shown in Figure 1. At the back there was a technician monitoring the functioning of the recording equipment and a medical doctor monitoring the EEG.

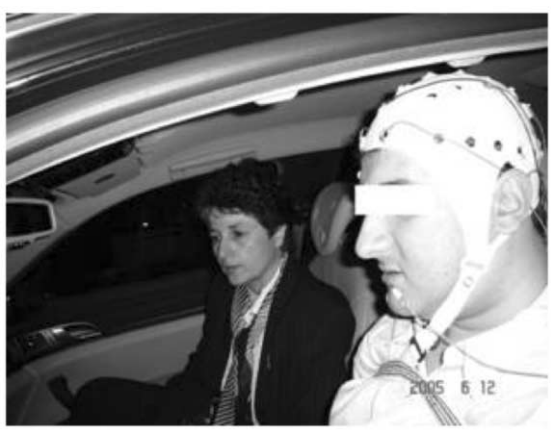


Figure 1 – The experimental car with the subject and the driving instructor

Every subject drove the research vehicle for a maximum of 1 hour on a motorway. In eight cases, subjects' sleepiness level during driving was very high, and the driver instructor stopped the measurements after three or more consecutive sleepiness events (unintentionally cross the lane border). The traffic on the motorway was very low, and the task was monotonous enough to stimulate hypovigilance. The driver's behavior and the sleepiness events for each subject were manually recorded. In parallel, the experimental car was equipped with the Eyelid Sensor System (ELS), i.e. a camera and appropriate commercial software capable of sensing the eyelid movements and detecting periods of drowsiness and sleepiness. In Table 1 there is an example of such a series of driving behavior annotations.

Table 1 – Example of annotations of the driver's condition and behavior

sec	sample	code	Notes
1170	234000	–1	ELS sleepy
1890	378000	–2	Theta waves in EEG
1895	379000	–3	ELS and EEG drowsy and sleepy
2575	515000	1	driver's error, wrong breaking
3185	637000	2	driver's error, serious driving event, wrong deviation from straight line, instructor's intervention
3360	672000	2	wrong deviation from straight line, instructor's intervention
3545	709000	2	instructor's intervention
3705	741000	3	end of measurement

Preprocessing

The EEG data were firstly filtered off-line by using a 3rd order Butterworth filter (band pass range: 0.5-45 Hz), and then the Infomax Independent Component Analysis (I-ICA) technique was used in order to remove eye movements and eye blinks [4]. ICA decomposition was performed on the EEG+EOG signals by using EEGLAB software [5]. Components contaminated by artifacts were rejected, and the remaining components were mixed and projected back onto the scalp channels. The analysis was performed on the artifacts-free EEG data.

Non-linear EEG analysis

A preliminary analysis took place for five subjects with various initial KSS scores. EEG channels C3, C4, P3 and P4 have been considered and analysis has taken place in time windows corresponding to 10 secs, with 25% overlap. Within each such EEG data segment the following features have been calculated:

- Sample entropy for the assessment of complexity of each channel, and

- Phase synchronization between two channels based on Hilbert Transform. Specifically C3-C4, C3-P3, C4-P4 and P3-P4 channel pairs have been considered.

The main idea behind this analysis was that both of these non-linear features can have relation to the systems dynamics and reveal changes during driver's gradual sleepiness and during incidents of driver's loss of control.

Sample entropy

A measure of the regularity of a time series is informative about the underlying complexities in the processes giving rise to it. Pincus [6] developed a regularity/complexity measure known as the Approximate Entropy (ApEn), which assigns a nonnegative number to a time series with larger values corresponding to greater apparent randomness of the process underlying the data while smaller values point to regular features in the data. Given groups of N points in a series, the approximate entropy is related to the probability that two sequences which are similar for N points remain similar at the next point.

A modification of the Approximate Entropy has been introduced as Sample Entropy[1] by Richman and Moorman [7], who claim that approximate entropy as defined by Pincus is biased. Larger Sample Entropy values indicate greater independence, less predictability, hence greater complexity in the data. This, in turn, may imply that decreased complexity or greater regularity in the time series is associated with disease or not regular function.

Considering two sequences $y_m(i)$ and $y_m(j)$ of the form $y_m(i)=[x(i), x(i+1),\ldots,x(i+(m-1))]$, then $B^m(r)$ gives the probability that the two sequences match for m points and $A^m(r)$ gives the probability of match for m+1 points and Sample Entropy is defined according to Equation 1.

$$SampEn(m,n,r) = [-\ln(\frac{A^m}{B^m})] \quad (1)$$

Phase synchronisation

Besides complexity measures, spatial features related to the dependence among EEG channels can be considered. Another approach is to consider the time-series as the output of oscillators and quantify their interaction by measuring the phase difference between these signals. In the present work, the analytic signal approach based on Hilbert transform has been adopted [8]. Hilbert transform can be defined as shown in Equation (2), where p.v. denotes the Cauchy principal value integral (which avoids the singularities at = t, and =±∞).

$$\tilde{s}(t) = \frac{1}{\pi} p.v. \int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau} d\tau \quad (2)$$

By use of Hilbert transform, the analytic $s_A(t)$ signal is defined from the initial signal s(t), as depicted in Equation (3)

$$s_A(t) = s(t) + j\tilde{s}(t) \equiv A(t)e^{j\varphi(t)} \quad (3)$$

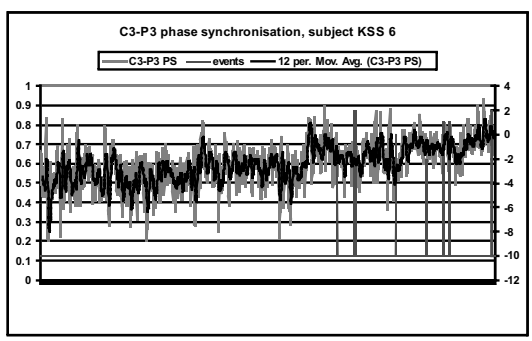
The instantaneous phase of the analytic signal is defined as shown in Equation (4)

$$\varphi(t) = \arctan\frac{\tilde{s}(t)}{s(t)} \quad (4)$$

Since instantaneous phase is available, the phase difference between channels can be calculated for each time sample. The measure of phase synchronization (PS) between two channels i and j within a time window of N samples is formulated as shown in Equation (5)

$$PS_{ij} = \frac{1}{N}\sum_{n=1}^{N} e^{j(\varphi_i(n)-\varphi_j(n))} \quad (5)$$

This is a normalized measure can vary between 0 and 1. Two channels which are synchronized for some time period will have X index close to 1 for that time window.



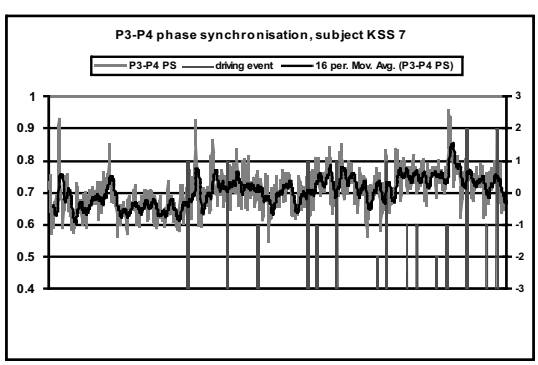


Figure 2 - Phase Synchronisation for two subjects. Synchronisation tends to increase. Events are related to periods of increased synchronization. The annotations follow the codes depicted in Table 1.

1 Matlab code for calculating Sample Entropy available at http://www.physionet.org/physiotools/sampen/

Results

The non-linear measures described above have been calculated for the five subjects with different initial sleepiness status. Since a series of such entropy and synchronization values was calculated, it was intended to seek for intra-subject differences in these features between the start and the end of the experiment, or between periods when a driving incident happened, or heavy sleepiness was reported and the rest of driving time.

Figure 2 shows the evolution of phase synchronization index and Figure 3 the evolution of Sample entropy for two subjecta, along with the annotated eventa, following the code depicted in Table 1. It can be observed that phase synchronization among the hemispheres is gradually increasing and sample entropy is decreasing, both pointing at a gradual increase of sleepiness related to loss of EEG complexity. Driving incidents seem to occur in relation to a local increase in phase synchronization and decrease in sample entropy.

In order to further evaluate possible variations of these features reflecting the subject's different condition, statistical analysis was performed. Specifically, different sets of feature data were considered, corresponding to: a) serious driving event, b) driving event, c) sleepy, d) drowsy areas, e) initial driving period without incidents and f) final driving period without incidents. The differences among some pairs of sets, i.e. serious driving event vs. initial driving period, were statistically assessed. The Wilcoxon ranksum test was applied, testing the hypothesis that populations generating two independent samples, x and y, are identical. Some results regarding these differences are depicted in Figure 4(a).

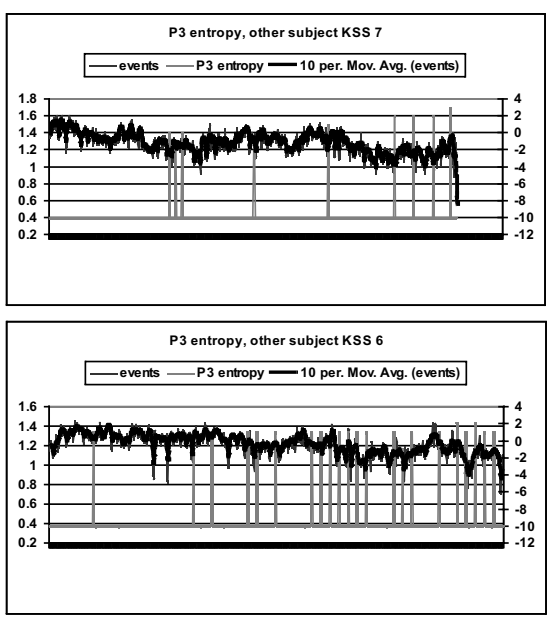


Figure 3- Sample entropy for two subjects. Events are related to periods of decreased entropy, and thus decreased complexity.

The statistical analysis showed that when there was statistical significance in the comparison, sample entropy for the aforementioned EEG channels was always lower during driving incidents than during normal driving periods. Moreover, it is interesting to notice a gradual change of the sample entropy group median values among the normal-sleepy-drowsy-event groups, as shown in Figure 4(b).

When the whole broadband signals were used in the analysis, a statistical difference in the PS feature between the normal driving period and the event period occurred for 4 out of 5 subjects. However, the expected under our hypothesis phase synchronisation elevation was evident for some cases, but not in all subjects. Specifically, an elevation clearly occurred in three out of five subjects, a decrease in two. For the sleepy/drowsy periods, the results were similar. It is possible that inter-subject variation is also related to the difference in sleepiness condition, as quantified by the KSS test. For example, the least number of statistically significant differences was found in the KSS 5 subject.

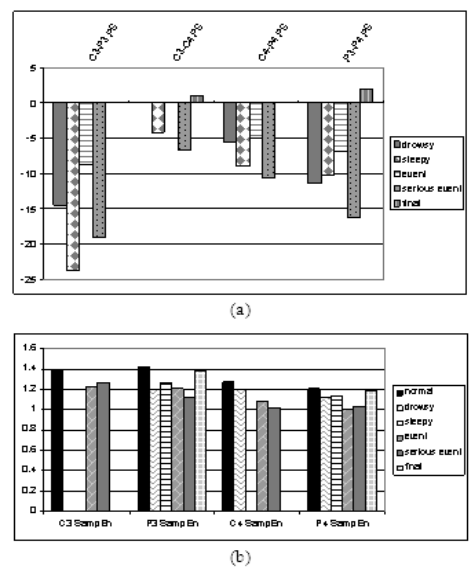


Figure 4 - a) For a subject KSS 7, the relative difference (1-value/initial_value)% between medians of the sample entropy calculated in the initial period and in the periods of various annotations. Calculations take place only when there is statistical significance. b) For another KSS 7 subject, median sample entropies for different channels and event periods than have statistical difference with the initial driving period.

When analyzing synchronization in different frequency bands (alpha and theta bands), there were some small differences, which were expected since more narrowband signals were analysed. In theta band, the statistically significant features showed a synchronisation increase for the driving event periods (4 out of 5 subjects), but that was not so persistent for the sleepy and drowsy periods. In alpha band, a phase synchronisation increase occurred for 3 out

of 5 subjects in the driving event and sleepy/drowsy periods.

It has to be noted that driver's loss of control and other annotated events do not occur momentarily, i.e. they refer to a period of time, for example not a specific second but a few seconds. Furthermore, they are manually reported by the co-driver, and their exact duration is unknown. Therefore, some inter-subject variation in the results may be explained by the variation in the definition and exact location of events. The accurate time annotation of driving events would allow the analysis or differences in pre and post event periods.

Discussion

An experiment with sleepy drivers has taken place and preliminary results of EGG data non-linear analysis have been presented. The decrease in Sample entropy and increase in Phase Synchronisation reported in this work can be regarded as measures of the system's complexity and spatial organization, complementary to the frequency band measures, traditionally applied in the sleep analysis. While these findings are somewhat macroscopic, a more extended study can seek more detailed patterns of sleepiness, to be used for prediction. A better elaboration of the driving event annotations would then be necessary. The correlation of polysomnographic signals or derived features with automatically generated driving data, as produced by the vehicle logging mechanism would be of use. This analysis can lead to an understanding of the mechanisms of alertness of drowsiness, and eventually to the prediction of dangerous driving events. The incorporation of such alertness predicting facilities in the cars can be extremely important for the driver safety. Such sensing platforms, as the one aimed by SENSATION project, must be real-time, autonomous, non-supervised, unobtrusive, and must be able to operate in non-constrained environment, able to monitor, detect and predict human physiological state in relation to alertness, fatigue and stress anytime, everywhere and for everybody.

Conclusions

A study of the intra-subject variation in sample entropy and phase synchronization in relation to driving accidents has been presented here. The non-linear measures, related to system's complexity and spatial organisation, can reveal overall system changes related to subject's vigilance, sleepiness, and hopefully predict conditions such as loss of control. The initial steady values can be used as reference, eliminating the needs for absolute thresholds. The elaboration of more detailed biosignal patterns related to the subject's driving behavior and the extended validation of our methodologies are among the future tasks of this research work.

Acknowledgments

This work has been partly supported by the IST - 2003- 507231 project SENSATION.

References

[1] Milter MM, Carskadon MA, Czeiler CZ, et al. Catastrophes. sleep and public policy: Consensus report. Sleep 1988: 11:100-09.

[2] Horne JA., Reyner LA., Sleep related vehicle accidents, BMJ 1995: 310(6979):565-7.

[3] Papadelis C, Kourtidou-Papadeli Ch, Bamidis P, Chouvarda I, Koufogiannis D, Bekiaris E, and Maglaveras N. Indicators of Sleepiness in an Ambulatory EEG Study of Night Driving. 28th IEEE EMBS Annual International Conference, Aug 30-Sept. 3, 2006, New York City, New York, USA.

[4] Jung TP, Makeig S, Humphries C, Lee TW, Mckeown MJ, Iraqui V and Sejnowski TJ. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 2000: 37: 163-178.

[5] Delorme A and Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 2004: 134: 9–21.

[6] Pincus SM. Approximate entropy as a measure of system complexity. Proc Natl Acad Sci USA 1991: 88: 2297-2301.

[7] Richman JS and Moorman JR Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 2000: 278(6): H2039-H2049.

[8] Rosenblum, M, Pikovsky A, and Kurths, J. Phase synchronization of chaotic oscillators. Phys. Rev. Lett. 1996: 76: 1804–1807.

Address for correspondence

Nicos Maglaveras
E-mail address nicmag@rned.auth.gr

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Identification and Genotype Related Classification of Children with Long QT-Syndrome Using 24h Holter Recordings

Matthias Bauch[a], Annamaria Siegler[a], Markus Khalil[b], Jörg Zehelein[c], Herbert E Ulmer[b], Hartmut Dickhaus[a]

[a] *Department of Medical Informatics, Heidelberg University, Germany*
[b] *Department of Paediatric Cardiology, University Medical Centre, Heidelberg, Germany*
[c] *Department of Internal Medicine III Cardiology, Angiology and Pneumology, University Medical Centre, Heidelberg, Germany*

Abstract

The long QT syndrome (LQTS) belongs to the family of hereditary diseases and can cause life-threatening arrhythmias and leads to sudden cardiac death. Mutations on six genes are responsible for changes in the electrophysiological properties of myocardial cells that are involved in the repolarization phase. In the surface ecg this is expressed by a prolonged QT interval and genotype-specific shapes for the T-Wave. The aim of the study was to find parameters that quantify properties of the repolarization phase which can be used in addition to the established Schwartz score in the process of diagnosing LQTS. Furthermore, ecg features were evaluated for the separation of the LQT subtypes LQT1, LQT2 and LQT3. The combination of the features PtA50 and QTc yielded with 93% sensitivity and 100% specificity the best results in the field of patient identification. Despite the small dataset consisting of 14 patients that was available for the second aim, the achieved results for the morphology indices motivate further research in this field.

Keywords:

Electrocardiography, Holter, long QT-syndrome.

Introduction

The long QT syndrome (LQTS) [1] belongs to the family of hereditary diseases which can cause life-threatening arrhythmias. Most of the patients suffer from first cardiac events like syncope or cardiac arrest during childhood. Approximately 54% of them die before they reach the age of twenty if no appropriate therapy was applied.

In clinical routine the diagnosis of the LQTS is essentially based on two aspects: information from the medical history and the evaluation of the ecg. The probability for the existence of a LQTS is given by a score proposed by Schwartz et al. in 1985 and 1993 [2]. The score is expressed by a number of points that are given for a prolonged QT interval, furthermore the former presence of Torsade de pointe arrhythmias, T wave alternans, notched T waves in at least three leads in the ecg, a lowered heart rate, former syncopes, congenital deafness, the fact that other family members have a LQTS and the occurrence of sudden death in the family. In addition to the Schwartz score the molecular genetic testing becomes more and more important. But this examination method is not suitable for screening tests because of the considerable costs and time. Nevertheless this is the only way to get evidence for the existence of a LQTS.

Particularly when applied to young children the Schwartz score shows some drawbacks. Especially for young patients most points that are related to the medical history are not achievable. Furthermore the widely used frequency correction by Bazett's formula often leads to unusable values for QTc if the original heart rate is very high which is a common property for young children. So a bigger part of the possible aspects in the Schwartz score are only partially suitable for the diagnostics of affected children. Therefore additional parameters are desired to identify patients correctly.

The LQTS encompasses a disease pattern that is based on mutations on six genes (KCNQ1 → LQT1, HERG → LQT2, SCN5A → LQT3, ANK2-ß → LQT4, KCNE1 → LQT5 and KCNE2 → LQT6) which cause functional changes in selected ion channels of the myocardial cells. This leads to prolonged action potentials of the cells and is represented by a longer QT interval in the surface ecg. Furthermore the changes on the different genes are responsible for different shapes of the repolarization phase in a heart cycle as various studies have shown [3-5].

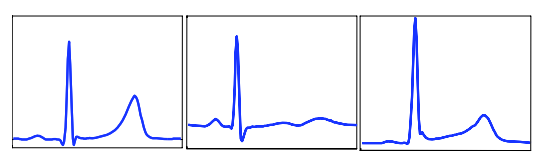


Figure 1 - Characteristic T wave shapes for LQT1 (left), LQT2 (middle) and LQT3 (right)

The drawbacks of the Schwartz score and the described T wave shape for the LQTS subtypes brings up the following

question: Do ecg parameters other than those mentioned in the Schwartz score improve the diagnostic process of LQTS? Furthermore it would be desirable that these features would also support the genotype based classification of identified patients. These two aspects might help in the diagnostic process of LQTS patients which are not identified by the Schwartz score. And additionally the time consuming process of a molecular genetic testing could be shortened if a probability for an affected gene is known.

Materials and methods

Materials and data preprocessing

24h Holter recordings from 14 genetically identified children with LQTS (age: 11,01±7,30 years; gender: f=7, m=7) and 22 healthy control persons (age: 11,31±3,75 years; gender: f=7, m=15) were used for the analysis. The group of patients is divided into three subgroups according to the genotype: LQT1 (9), LQT2 (3) and LQT3 (2).

The ecg recordings were performed using digital twelve channels recorders from Mortara Instrument Inc (Milwaukee, USA). In a first stage the recordings were processed using the Mortara Holter software H-Scribe for the identification and annotation of normal beats. All further processing was carried out using a self developed software framework for biosignal analysis [6]. An ecg viewer which is part of the framework was used to manually add missing beats that were not detected by the Holter software. Superimposed noise was removed using a highpass filter with a cut off frequency of 0,5 Hz and a lowpass filter with a cut off frequency of 50 Hz. Baseline wander was corrected using a cubic spline interpolation.

All further processing of the data except for the morphological indices was carried out using only lead V5. Due to the most distinct depiction of the repolarization phase lead V5 was chosen. Another reason is that lead V5 has the best signal to noise ratio in the considered population. To take daytime dependencies into account, three sections of the recordings were analyzed. These segments were taken from a time window in the morning starting at 8 am, from the evening starting at 6 pm and at night starting at 1 am. From these segments sequences of 10, 20, 50 and 100 consecutive normal beats are used for the analysis.

In all selected segments a self developed delineation routine to find the correct positions of Q, R, R_{Peak}, S, J, T_{Peak} and T_{End} was applied [7].

Methods

Because of the importance of the repolarization phase for the LQTS particular T wave parameters were quantified which can be assigned to the categories time interval measurements, area measurements, and morphology indices.

Time interval measurements

This category contains the features that are widely used in clinical routine. These are the intervals between prominent marker positions which were automatically inserted during the delineation process. Figure 2 demonstrates the considered intervals. All measures of this category were calculated as absolute and frequency corrected values using Bazett's formula.

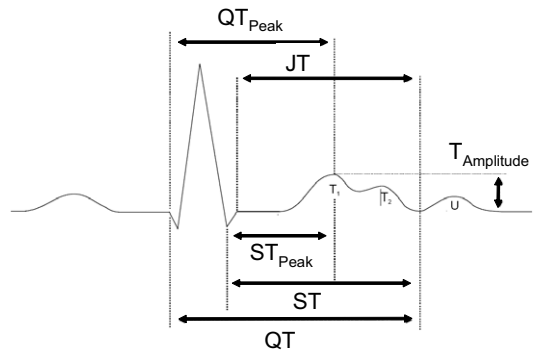


Figure 2 - Time interval measurements

Area measurements

The following area-derived parameters contain also information about the repolarization phase [8]. Table 1 summarizes the parameters in the area measurement category.

Table 1 - Area measurements

Feature	Description
A_{tot}	Total area of the repolarization phase
t_{A50}	Time interval to accumulate 50% of the area.
Pt_{A50}	Percentage of repolarization time needed to accumulate the first 50% of the area.
$t_{A25\text{-}50}$	Time interval to accumulate the mid 50% of the repolarization area from 25% to 75%.
Pt_{A25_75}	Percentage of repolarization time needed to accumulate the mid 50% of the repolarization area from 25% to 75%.
t_{A97}	Time interval to accumulate 97% of the total repolarization area
T_{amp}	Maximum amplitude of the T wave

Morphology indices

The calculated morphology indices are applied as proposed by Acar et al [9]. They characterize the spatial and temporal variations of the T wave morphology and wavefront direction differences between the depolarization and repolarization phase.

In contrast to the already described parameters the morphology indices are applied to a vector curve. It is constructed by applying a singular value decomposition on the data matrix whose rows correspond to the ecg leads I,

II, V1-V6. A summary of the morphology indices is given in table 2.

Table 2 - Morphology indices

Feature	Description
TMD	T wave morphology dispersion. Represents the variation of the morphology of the repolarization phase between different leads
TMD_{pre}	The same as TMD, but takes only the rising edge of the T wave into account
TMD_{post}	The same as TMD_{pre} with respect to the falling edge of the T wave
TCRT	Total cosine R to T. Quantifies the angle between the R- and T-wave loop
PL	Percentage of the inner area in a 2D projection that is covered by the T wave
PO	Percentage of the outer area in a 2D projection that is covered by the T wave
LD_1, LD_2	Lead dispersion 1 and 2: variation of the ecg vector during the repolarization phase

Statistical evaluation

The discriminative power of single parameters was evaluated using receiver operating characteristics (ROC) plots. Combinations of up to three parameters were used as inputs for a second order polynomial classifier. The selectivity of all results is given by values of sensitivity and specificity.

Results

As mentioned above all features were calculated for sequences of consecutive normal beats at specific times of the day. By means of this approach time dependent effects should be considered. All parameters were calculated for each single beat in the corresponding segments. To aggregate the results mean values for the parameters were calculated for the first 10, 20, 50 and 100 consecutive beats.

Patient identification

The best results for single features are given by the time interval measurements with rates of correct classified datasets of up to 97%. For two reasons these results were excluded in the further investigations. First of all the widely used Bazett formula for the frequency correction yields erroneous data sets if the heart rate is too high which is very common in young childrens. The second reason for the exclusion is the strong correlation between the parameters in this category and the fact that QTc is already part of the Schwartz score. For the remaining features those from the category of area measurements yield the best results. Table 3 lists the best features to discriminate the LQTS patients from the healthy subjects.

As table 3 shows, the best separation of the two groups is achieved in the evening segment using the parameter Pt_{A50} with values of 86% for the sensitivity and 100% for the specificity. This yields a rate of correct classifications of 93%.

Table 3 - Best single features for the identification of LQTS patients. The discriminative power is given by values for sensitivity and specificity in percentage

	10 Beats	20 Beats	50 Beats	100 Beats
Morning 09 am	t_{A50}: 86 / 91	t_{A50}: 79 / 91	t_{A50}: 79 / 91	t_{A50}: 79 / 95
Evening 06 pm	Pt_{A50}: 86 / 95	Pt_{A50}: 86 / 100	Pt_{A50}: 86 / 100	Pt_{A50}: 86 / 95
Night 01 am	Pt_{A50}: 71 / 86	t_{A25_75}: 64 / 91	t_{A25_75}: 64 / 95	Pt_{A50}: 64 / 91

The results in table 3 also indicate that the number of feature values used in the averaging process is not significant. Reliable values were achieved with a minimum number of 20 beats. The results for the morphological parameters are not suitable to separate the two groups of the study.

Combinations of two features were also evaluated. The criteria for parameters to be used in this task is given by values for the sensitivity and specificity greater than 70%. Table 4 lists the best combinations of two parameters.

Table 4 - Best feature combinations for the identification of LQTS patients. The discriminative power is given by values for sensitivity and specificity in percentage

	10 Beats	20 Beats	50 Beats	100 Beats
	2 Features			
Morning 09 am	Pt_{A50}, QTc: 86 / 100	Pt_{A50}, SPTc: 86 / 100	Pt_{A50}, QTc: 86 / 100	Pt_{A50}, QTPc: 86 / 100
Evening 06 pm	Pt_{A50}, SPTP: 86 / 100	Pt_{A50}, QTc: 86 / 100	Pt_{A50}, QTc: 93 / 100	Pt_{A50}, QTc: 93 / 100
Night 01 am	Pt_{A50}, SPTP: 93 / 91	Pt_{A50}, SPTP: 93 / 91	Pt_{A50}, SPTP: 93 / 91	t_{A25_75}, QTPc: 71 / 100

As for single features the best results for parameter combinations are achieved in the evening. A rate of 97% correctly classified datasets was achieved for the combination of Pt_{A50} and QTc.

Genotype related classification

Due to the small number of patients in the three target groups LQT1 (9 children), LQT2 (2 children) and LQT3 (3 children) a reliable classification was not possible. Nevertheless single features were evaluated with respect to their possible application for this task.

The most promising results were achieved by the area measurement t_{A50} and the morphology index TCRT. For t_{A50} the rate of overall correctly classified patients was 79% whereas TCRT yielded 77%. Despite the small sample size these result can be used as a first indicator for a genotype related classification of LQTS patients.

Discussion

The established gold standard in the diagnosis of the Long QT-syndrome is the Schwartz score as mentioned in the introduction. But especially in the diagnosis of affected young children and babies this approach often gives non-optimal results due to two limiting factors. The first one is that the frequency correction according to the Bazett formula gives erroneous values for recordings with high frequencies. Second the aspects from the previous medical history can not be taken into account as they are often not available for the young patients. Therefore the accomplished scores are often to low for the identification of affected patients. As mentioned before our work addresses these deficits by adding additional information derived from the analysis of ecg recording to the diagnostic process. These new features are not intended to replace the Schwartz score but to be used in addition especially when applied to children.

Besides of the QT interval the prolongation of the T wave can be clearly quantified by the area derived measurements as our results indicate. Approximately 68% of the whole time for the repolarization phase is needed to fill the first 50% of the covered area (Pt_{A50}) whereas healthy test persons needed only 62% of the time.

Interesting is also the fact that the used morphology indices are not suitable to describe the characteristic T wave patterns that were qualitatively described in other studies [3-5].

As our findings indicate, the area measurements and partly the morphology indices are suitable to identify LQTS patients. In combination with the traditional time interval parameters they yield the best results in the classification process. This is also in accordance with the findings of Struijk and co-workers [10] who did similar investigations on resting ecg recordings obtained from adults.

As is known, time series constructed from RR or QT intervals exhibit circadian properties. This issue was targeted during our work by the examination of three segments from the recodings covering the morning, the evening and the night. But our results indicate that circadian effects have only little influence on the discriminative power of the used features as table 3 and 5 demonstrate.

Although our findings indicate good separative properties for the time interval, area and morphology measurements they have to be interpreted with care as they are based on a relatively small dataset. The reason for this small dataset is the fact that although most of the identified patients underwent genetical testing no established mutations were found on the genes.

The results achieved so far and the aims to find ecg features that can be used in addition to the Schwartz score future investigations are needed. In this study all patients had definite results for the Schwartz score. But in cases with borderline or negative results the experience of the investigating physician is needed. So the next task is to analyze recordings from this special group of patients to validate our findings.

Conclusion

Using a collection of parameters from the categories of time interval, area and morphology measurements a good separation of patients with LQTS from healthy control persons could be achieved. The parameters were selected complementary to those mentioned in the Schwartz score and introduce new aspects in the process of diagnosing LQTS. Future work should focus especially on the question whether patients who are not identified by the Schwartz score can be recognized using the additional ecg features. Furthermore ecg parameters were evaluated that might be used to discriminate the most common LQT subtypes LQT1, LQT2 and LQT3. Although only a small dataset was available the achieved results motivate for further investigations.

Acknowledgments

This work is supported by a grant from the Ministry of Science, Research and the Arts Baden-Württemberg, Germany

References

[1] Schwartz PJ, Priori SG and Napolitano C. The Long QT Syndrome. In: Zipes DP and Jalife J. eds. Cardiac Electrophysiology – From Cell to Bedside. W.B. Saunders Company, 2000; pp 597-615

[2] Schwartz PJ, Moss AJ, Vincent GM, Crampton RS. Diagnostic criteria for the long QT syndrome. An update; Circulation. 1993 Aug; 88(2):782-4.

[3] Zareba W. Genotype-specific ECG patterns in long QT syndrome. J Electrocardiol. 39 (2006) S101-6

[4] Moss AJ, Zareba W, Benhorin J, Locati EH, Hall WJ, Robinson JL, Schwartz PJ, Towbin JA, Vincent GM and Lehmann MH. ECG T-wave patterns in genetically distinct forms of the hereditary long QT syndrome. Circulation. 1995; 92:2929-34.

[5] Zhang L, Timothy KW, Vincent GM, Lehmann MH, Fox J, Giuli LC, Shen J, Splawski I, Priori SG, Compton SJ, Yanowitz F, Benhorin J, Moss AJ, Schwartz PJ, Robinson JL, Wang Q, Zareba W, Keating MT, Towbin JA, Napolitano C and Medina A. Spectrum of ST-T-wave patterns and repolarization parameters in congenital long-QT syndrome: ECG findings identify genotypes. Circulation. 2000;102:2849-55.

[6] Dickhaus H, Maier C, Bauch M, Rengshausen U, Khalil M, Hessling G and Ulmer HE. A System for Automated Analysis of RR-Series Applied to Identification of Patients with Neurocardiogenic Syncope, Proceedings of the Second Joint EMBS/BMES Conference Houston, TX, USA, 2002, 1620-1.

[7] Siegler A. Genotyp-Phänotyp Relation bei Patienten mit Long QT-Syndrom durch das 24-Stunden EKG. Master Thesis. Heilbronn University 2006.

[8] Zareba W, Moss AJ and Konecki J. TU Wave Area-Derived Measures of Repolarization Dispersion in the Long QT Syndrome. J Electrocardiol. 1998; 30 Suppl:191-5.

[9] Acar B, Yi G, Hnatkova K and Malik M. Spatial, temporal and wavefront direction characteristics of 12-lead T wave morphology. Med Biol Eng Comput. 1999; 37(5): 574–584.

[10] Struijk JJ, Kanters JK, Andersen MP, Hardahl T, Graff C, Christiansen M and Toft E. Classification of the long QT syndrome based on discriminant analysis of T-wave morphology. Computers in Cardiology. 2005; 25-28: 511- 514

Address for correspondence

Dipl.-Inform. Med. Matthias Bauch
Department of Medical Informatics
Heidelberg University
Im Neuenheimer Feld 400
D-69120 Heidelberg
Email: mbauch@hs-heilbronn.de

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Temporal Abstraction and Data Mining with Visualization of Laboratory Data

Katsuhiko Takabayashi [a], Tu Bao Ho [b], Hideto Yokoi [c], Trong Dung Nguyen [b], Saori Kawasaki [b], Si Quang Le [b], Takahiro Suzuki [a], Osamu Yokosuka [d],

[a] *Division for Medical Informatics and Management, Chiba University Hospital, Inohana, Chuou-ku, Chiba, 260-8677 Japan*
[b] *Japan Advanced Institute of Science and Technology,Tatsunokuchi, Ishikawa, 923-1292 Japan*
[c] *Division for Medical Informatics, Kagawa University Hospital, Ikedo, Miki, Kida-gun, Kagawa, 761-0793 Japan*
[d] *Department of Gastroenteology, Chiba University Hospital, Inohana, Chuo-ku, Chiba, 260-8677 Japan*

Abstract

To analyze the laboratory data by data mining, user-centered universal tools have not been available in medicine. We analyzed 1,565,877 laboratory data of 771 patients with viral hepatitis in order to find the difference of the temporal changes in laboratory test data between Hepatitis B and Hepatitis C by the combination of temporal abstraction and data mining. The data for one patient is temporal for more than 5 years. After pretreatment the data was converted to abstract patterns and then selected into sets of data combination and rules to identify Hepatitis B or C by D2MS and LUPC which were originally produced by ourselves. Not only data pattern, but also temporal relations were considered as a part of the rules. In the course of evaluating the results by domain experts, even though there were not so remarkable hypotheses, visualization tools made it easier for them to understand the relations of the complicated rules.

Keywords:

databases, liver function test, hepatitis

Introduction

Data mining is beginning in medical application. Even though there are many data mining techniques to analyze the database, most of them are still experimental and in the hands of computer scientists. Medical doctors cannot use the tools for their own data analysis individually. User-centered universal tools should be applied for medical researchers to analyze their own data. In Japan we started the national project of active mining to analyze the same database of viral hepatitis and evaluate the availability and validity to implement it in medicine[1]. In this approach from the medical point of view, not only providing data, but we performed pretreatment of medical data and attended evaluation with mathematical researchers as active mining. One of the keys of laboratory data analysis in medicine is how to treat temporal data. By using temporal abstraction, we aimed to solve this problem.

The goals

One of the concrete goals of this project was to discover the differences in the temporal patterns of hepatitis B (HBV)and C (HCV) which has not been clearly defined, and, more importantly, to examine whether the methods we applied here can work well and be applied to other fields.

Materials and methods

Database

The hepatitis data is located in two databases, one is the laboratory database of the hospital, and another is the biopsy and clinical data in the department of Medical School, Chiba University. The contents are as follows:

- Basic information of patients (total 771 records)
- Results of biopsy (total 960 records)
- Information about measurements in in-hospital tests (total 459 records)
- Results of out-hospital tests (total 30,243 records) and those of in-hospital tests (total 1,565,877 records)

The data in the hospital are more than 20 years. Those patients were performed liver biopsy once at least.

Preprocessing

Our preprocessing of hepatitis data includes data cleaning, data integration, data reduction, deidentification and data transformation. We removed only redundant and not unsuitable suffix data for further processing and we eliminate noisy data in the next temporal abstraction step. For the purpose of temporal abstraction, we have to integrate original relational data tables into one data table, where each column represents laboratory examination. By combining the expert guidance and the frequencies of attributes presented from 983 examinations, we selected 15 most significant examinations. The numbers of examinations for each patient are different, and the examination periods are irregular.

Temporal abstraction

Temporal abstraction (TA) is one approach to deal with time-related data in medical research. TA methods are

those able to derive an abstract description of temporal data by

extracting their most relevant features[2] .The key idea is to transform time-stamp points by abstraction into an interval-based representation of data by extracting their most relevant features [3]. TA task can be defined as follows.

The input includes a set of time-stamped data points (events). The output includes a set of interval-based, context-specific unified values or patterns (usually qualitative) at a higher level of abstraction. TA can be generally considered in two phases: *basic* TA for abstracting time-stamped data from given episodes (which are significant intervals for the investigation purpose) and *complex* TA for investigating specific temporal relationships between episodes that can be generated from a basic TA or from other complex TAs.

Basic temporal relations

We started by a separation of two groups of tests, one with values that change rapidly in the short term such as GOT, GPT, TTT and ZTT and the other with values that change slowly in the long term such as T-CHO, CHE, ALB.

Basic temporal abstractions typically extract *states* (e.g., low, normal, high), and/or *trends* (e.g., increase, stable, decrease) from a uni-dimensional temporal sequence.

The essential ideas of our temporal abstraction methods here is to deal with long and irregular time-stamp sequences, and doing abstraction in efficient. We introduce the notion of "changes of state" to characterize the slowly changing tests, and the notions of "base state" and "peaks" to characterize the rapidly changing tests .

Temporal abstraction primitives

From observation and analysis, we defined the following temporal abstraction primitives:

1. State primitives: N (normal), L (low), VL (very low), XL (extreme low), H (high), VH (very high), XH (extreme high).
2. Trend primitives: S (stable), I (increasing), FI (fast increasing), D (decreasing), and FD (fast decreasing).
3. Peak primitives: P (peaks occurred).
4. Relations: ">" (change state to), "&" (and), "–" (and then), "/" ("X/Y" means majority of points are in state X and minority of points are in state Y).

The thresholds to distinguish the state primitives of tests are given by medical doctors, for example, those of the test GOT are 40, 100, 200, respectively. We define structures of abstraction patterns as follows:

<pattern> ::= <state primitive>

<pattern> ::= <state primitive> <relation>

<pattern> ::= <state primitive> <relation> <peak>

<pattern> ::= <state primitive> <relation> <state primitive>

Examples of abstracted patterns in a given episode are like follows:

GOT = H (all the values of GOT in one case are above the normal region as shown in the upper left in Figure 1),

"GPT = H&P" (The values of GPT in one case are always high and with peaks like lower left of Fig 1),

"I-BIL = N>L>N" (I-BIL first is normal, then changed to the low region, and finally changed to the normal region in one case like the right bottom in Fig 2) etc.

Figure 1 shows typical possible patterns (8 and undetermined) for rapidly changing tests, and *Figure 2* shows typical possible patterns (21 and undetermined) for slowly changing tests [4].

Abstraction of rapidly changing test results

From our observation and analysis, especially GPT and GOT were defined as rapidly changing attributes, which can go up in a very short period and go back to a stable" state. Thus two most representative characteristics of these tests are a "stable" base state (BS), and the position and value of peaks, where the attributes suddenly go up. Based on this, we formulated the following algorithm to find the base state and peaks of a test. Rapidly changing tests applied also to TTT and ZTT and they showed 9 patterns.

Algorithm 1 (for rapidly changing tests)

Input: A sequence of patient's values of a test with length N denoted as S_{00} = {s1, s2, …, sN} in a given episode.

Output: Base state and peaks, and an abstraction of the sequence derived from them.

Parameters: NU, HU, VHU, XHU: upper thresholds of normal, high, very high, extreme high regions of a test, a (real).

Notation:

- Mi: Set of local maximum points of S
- BS: base state of S
- PEi: set of peaks of S

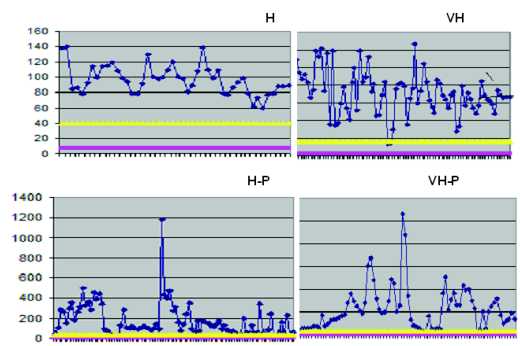


Figure 1 - rapidly changing test patterns

Abstraction of slowly changing test results

The key idea is to use the "change of state" as the main feature to characterize sequences of the tests. It can be seen that the "change of state" characterize information of both state and trend of the sequences.

From the beginning of a sequence, the first data points can be at one of the three states “N”, “H”, or “L”. It will happen that:

Either the sequence changes from one state to another state, smoothly or variably (at boundaries), or the sequence remains in its state without changing.

We provided 22patterns for slowly changing data.

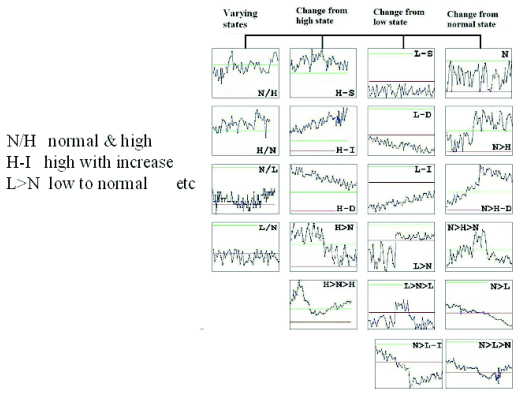


Figure 2 - Slowly changing test patterns

Temporal relationships

The temporal relations between the abstracted events of laboratory data were also treated here as phenomena and a part of the rules by comparing the period of the state. They are classified into seven relations by Allen[8].

$$Z = \frac{conf(R) - p(C)}{\sqrt{p(C)(1-p(C))/n_A}}$$

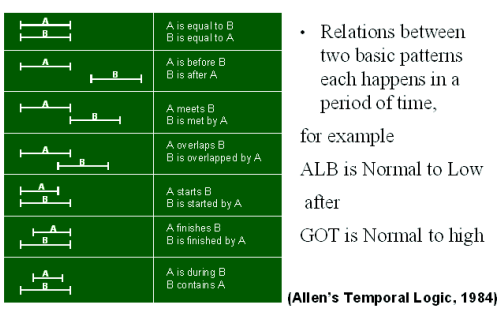


Figure 3 - Temporal relations

Complex temporal abstraction

Mining abstracted hepatitis data with system D2MS

The authors developed an interactive visualization tool in decision tree construction called D2MS (data mining for Model Selection) for supporting an effective cooperation of the user and the computer in classification. D2MS shares many features with WinViz [5] and Cviz that both use parallel coordinates. WinViz allows the user to visually examine a tabular database and to formulate query interactively and visually. Cviz is an attempt to integrate visualization into the knowledge discovery process.

D2MS facilitates the trials of various alternatives of algorithm combinations and their settings. The data mining methods in D2MS consists of programs CABRO for tree learning and LUPC for rule learning [6]. CABRO produces decision trees using R-measure and graphically represents them in particular with T2.5D tool (trees 2.5 dimension). As shown in *Figure 4*, visualization made us easily recognized the different pattern of HBV and HCV.

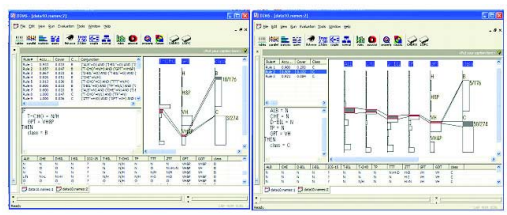


Figure 4 - Rules for HBV and HCV by D2MS

We examined statistical significance of the consequence according to the method of [7], which prunes discovered rules statistically as follows.

Assume a rule R: A C (or R: A ¬C) with confidence conf(R). If conf (R) = p(C) then R is eliminated. To test whether conf (R) = p(C), we use the following test statistic where n_A is the number of cases satisfying C

Mining abstracted hepatitis data with LUPC

LUPC is a separate-and-conquer algorithm that controls the induction process by several parameters. The parameters allow obtaining different results and this ability allows the user to play a central role in the mining process [6].

LUPC is developed to learn prediction rules from supervised data. Each rule found by LUPC is a conjunction of attribute-value pairs that may present an interesting pattern. The main features of LUPC are (1) its ability of finding rules with associate domain knowledge (such as finding rules containing or not containing specified attribute-value pairs), as well as finding rules for minority classes; (2) it is integrated with D2MS’s rule visualizer and thus supports the user in selecting the appropriate rules which result from different possible settings of parameters. The performance of LUPC depends on several parameter specified by the user: for min accuracy of rules, for min coverage of rules, for maximal number of candidate

rules in the beam search, and for maximal number of attribute-value pairs to be consider. By varying these parameters we can find different sets of rules [6]. When using the setting with default parameters of = 80%, = 3, = 200, and = 100, we found 119 rules characterizing the hepatitis B and 152 rules characterizing hepatitis C.

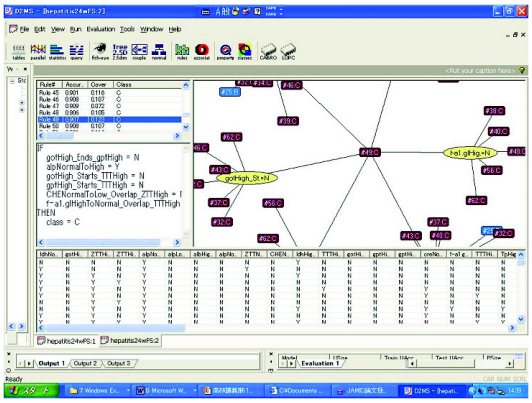


Figure 5 - LUPC Rules can be illustrated in a left figure

Evaluation

The produced rules were evaluated by three experts.

Results

By using D2MS, we discovered 33,477 rules for type B and C difference. These rules are complicated and sometimes contradictory to each other. For example, there is a rule if T-Bilirubin(Bil) is N(normal) , ZTT is N, and GOT is N with P(peak), then HBV, while there is another rule if D-Bil is N, TTT is N, GOT is N with P, then HCV, which are almost the same but the results are completely different.

After pruning by statistical aspects between HBV and HCV, there are only 27 rules (0.08%) left (*Table1*). However, these rules seemed not attractive for medical doctors even if they are statistically significant. For example, T-cholesterol is normal is HCV in 171/260 cases (66%), or GPT is high with a peak and ZTT is mostly high partly normal is HBV. They are too simple or vague and must be carefully assessed.

Different datasets were found by using LUPC with various parameters including temporal relations between laboratory tests. *Table2* presents the top five rules by LUPC from the point of coverage and confidence. From this table, especially rule3 and 5 are similar and could be merged as a rule that if TP decreasing high to normal and both ZTT and TTT is high, then it is HCV. In the evaluation of medical doctors, though most of them seemed not crucial or not useful in clinical medicine even the discovered rules covering many cases with high accuracy. However, some of the rules could be reasonable from the different clinical course of two types of hepatitis, especially when the experts checked and integrated the rules in the illustration.

Table 1 - 27 pruned rules produced by D2MS and satisfying chi-square test

No.	class	T-CHO	CHE	GOT	GPT	TTT	ZTT	D-BIL	T-BIL	I-BIL	TP	acc	ratio
1	C (95%)	N										0.66	171/260
2	C (95%)		N									0.72	183/256
3	C (95%)							N				0.73	180/248
4	C (95%)			H								0.76	89/117
5	C (95%)				H							0.76	78/103
6	C (95%)		N					N				0.82	142/173
7	C (95%)					N	H/N					0.92	11/12
8	C (95%)				H				N/H			0.93	14/15
9	B (95%)				H&P		H/N					0.92	11/12
10	B (90%)						N					0.68	63/93
11	B (90%)		H/N									0.7	14/20
12	B (90%)		N/H									0.74	23/31
13	B (90%)			N&P	H&P							0.7	16/23
14	B (90%)		N				N/H	N				0.88	7/8
15	C (90%)		N			N						0.8	67/84
16	C (90%)						H-I	N				0.95	63/66
17	C (90%)			H	XH							0.92	11/12
18	C (90%)		N	N					N			0.93	26/28
19	C (90%)										N/H	0.81	35/43
20	C (90%)					H-I				N		0.93	25/27
21	C (85%)				N&P	N						0.69	33/40
22	C (85%)		N	N								0.84	41/49
23	C (85%)		N	H								0.87	58/67
24	C (85%)					H/N				N		0.79	23/29
25	C (85%)				XH							0.72	18/25
26	C (85%)	N/L										0.8	28/35
27	C (85%)						H-D					0.83	33/40

Table 2 - Top 5 rules selected by LUPC

Rule 1(Coverage ; 4.098% confidence ; 100% coverage ; 25 cases)

creNormalToLow = Y — Creatinine decreasing from normal to low
gptHigh_Start_gotHigh = Y — and GPThigh start with GOT high
-> Class = C — is HCV

Rule 2(Coverage 3.443% confidence ;100% coverage;21 cases)
Bilirubin decreasing from normal to low before TTT elevates
is HCV

Rule 3 (Coverage 3.443% confidence ;100% coverage;21 cases)
TP decreasing high to normal and ZTT goes up to high after TTT
up to high
is HCV

Rule 4 (coverage 3.279% confidence 100% coverage; 20 cases)
Creatinine decresing normal to low and bilirubin increasing normal
to high
is HCV

Rule 5 (coverage 2.951% confidence 100 % coverage; 18 cases)
TP decreasing high to normal and ZTT goes up to high before TTT
up to high
is HCV

In the viewer of LUPC we can see the accuracy and coverage on the upper left, rule itself in the middle and the relation of the rules and attribute value pairs in the figure to the right which can be manipulated by users (*Figure 6*). By handling it users can see the relations of each item. In this figure doctors could easily recognize if LDH is low to normal is false, then all the rules are related to HCV, while if creatinine is normal to low is false, then it is related to the rules of HBV except rule #48(center), so that the doctors could understand from a more comprehensive point. This technique was highly evaluated by medical doctors and some rules such as the top 5 were considered as meaningful.

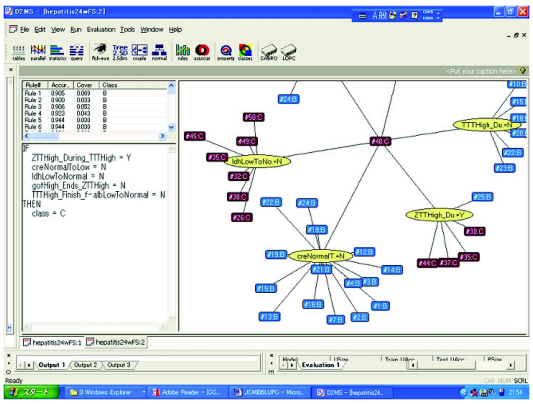


Figure 6 - LUPC makes it easier for the users to understand the relations comprehensively

Discussion

We have presented a temporal abstraction approach to data mining the temporal hepatitis data. Most doctors do not believe that they can distinguish HBV and HCV from the laboratory data change, so it might be true when we cannot obtain any new findings from this data mining. However, we in fact obtained many rules to identify HBV or HCV with statistical significance. Some of them look very simple. To confirm, we need to analyze them by stratified analysis, removing the modification of other factors such as treatment and then compare again. By LUPC, we can estimate the patient data comprehensively and expect new findings in many diseases because it is difficult for human beings to find data changes over a long period of time. some experts of liver diseases mentioned that the cases with HBV and HCV were apt to show different clinical changes, and our results would reflect these changes.

One of the major problems in rule based data mining is that there are too many rules deduced for us to evaluate. To select most important rules we introduced the chi-square test which was effective to decrease the number of rules as well as statistical reasoning. Another is by LUPC, not only selecting minority classes from large unbalanced datasets but visualization, it is not difficult to separate the important ones from many rules for medical doctors.

Other studies of data mining in medicine are mostly in the field of genomics and epidemiology and the analysis of laboratory data is quite limited. We provided the data of anti-phospholipid antibody syndrome to PKDD 1999 as model data in order to establish a new technique as well as hepatitis data.

Current medical study is deeply inclined to use a prospective way or Cohort as a scientific study design, which implies a carefully planed experiment. However, when we think of a long-term experiment lasting over 10 years, it is not realistic to study prospectively. There is a great possibility of new paradigms appearing before the study is completed. Retrospective studies are expected for these long term studies and data mining techniques will play a major role in this filed by creating high potential hypotheses.

Even though we did not discover crucial rules to show the difference of laboratory data change between HBV and HCV, we proved to show that this combination of TA and data mining with visualization is useful and effective. Furthermore, it could be applied to other fields of medicine and would be a basic model for the universal analysis of data mining for temporal data analysis in medicine.

Conclusion

The rules that show the difference of the laboratory changes in the long clinical course between the HBV and HCV could be deduced by D2MS. Pruning by statistical significance could decrease the number of rules but obtained rules were not interesting in individuals. Visualization made it easier for doctors to find the relations and led them to find reasonable results. This combination technique of temporal abstraction and data mining with visualization could be applied universally.

References

[1] Motoda, H., Active Mining: New Directions of Data Mining, IOS Press, 2002.

[2] Horn, W., Miksch, S., Egghart, G., Popow, C., Paky, F., "Effective Data Validation of High-Frequency Data: Time-Point-, Time-Interval-, and Trend-Based Methods", *Computer in Biology and Medicine, Special Issue:Time-Oriented Systems in Medicine*, 27(5), 389-409, 1997.

[3] Bellazzi, R., Larizza, C., Magni, P., Monntani, S., and Stefanelli,M. Intelligent Analysis of Clinic Time Series: An Application in the Diabetes Mellitus Domain *Artificial Intelligence in Medicine*, 20, 37-57, 2000.

[4] Ho, T.B., Nguyen, T.D., Kawasaki, S., Le, S.Q., Nguyen, D.D., Yokoi, H., Takabayashi, K. Mining Hepatitis Data with Temporal Abstraction. *ACM Inter. Conf. on Knowledge* Discovery and Data Mining KDD-03, 369-377,2003.

[5] Lee, H.Y., Ong, H.L., and Quek, L.H. Exploiting Visualization in Knowledge Discovery. *First Inter. Conf. on Knowledge Discovery and Data Mining,* 1995, pp. 198–203.

[6] Ho T.B., Nguyen T.D., Nguyen D.D., Kawasaki S., Visualization Support for User-Centered Model Selection in Knowledge Discovery and Data Mining. *Int. Journal Artificial Intelligence Tools,* 2001;10(4): 691-713.

[7] Bruzzese, D. and Davino, C. Statistical Pruning of Discovered Association Rules. *Computational Statistics,* 16 (3), pp. 387 –398, 2001.

[8] Allen E and Jai S. A decidable temporal logic to reason about many processes. Proceedings of the ninth annual ACM symposium on Principles of distributed computing table of contents. Quebec City , Canada 233-246,1990.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Automated Interpretation of Optic Nerve Images: A Data Mining Framework for Glaucoma Diagnostic Support

Syed SR. Abidi[a], Paul H. Artes[b], Sanjan Yun[a], Jin Yu[a]

[a] *NICHE Research Group, Faculty of Computer Science, Dalhousie University, Halifax, Canada*
[b] *Department of Ophthalmology and Visual Sciences, Dalhousie University, Halifax, Canada*

Abstract

Confocal Scanning Laser Tomography (CSLT) techniques capture high-quality images of the optic disc (the retinal region where the optic nerve exits the eye) that are used in the diagnosis and monitoring of glaucoma. We present a hybrid framework, combining image processing and data mining methods, to support the interpretation of CSLT optic nerve images. Our framework features (a) Zernike moment methods to derive shape information from optic disc images; (b) classification of optic disc images, based on shape information, to distinguish between healthy and glaucomatous optic discs. We apply Multi Layer Perceptrons, Support Vector Machines and Bayesian Networks for feature sub-set selection and image classification; and (c) clustering of optic disc images, based on shape information, using Self-Organizing Maps to visualize sub-types of glaucomatous optic disc damage. Our framework offers an automated and objective analysis of optic nerve images that can potentially support both diagnosis and monitoring of glaucoma.

Keywords:

glaucoma, data mining, feature selection, clustering, Confocal Scanning Laser Tomography, Support Vector Machines

Introduction

Glaucoma is an eye disease that is characterized by slow progressive damage to the optic disc and corresponding deterioration of the patient's vision [1]. At present, there is a gap in the understanding of the cause, the types and the natural course of glaucoma. The use of sophisticated imaging technologies, such as Confocal Scanning Laser Tomography (CSLT), capture 3-dimensional images of the optic disc that are used for diagnostic purposes [2]. However, the interpretation of CSLT images is a manual and subjective process—a trained professional has to manually define the margins of the optic nerve based on his/her training and expertise and then classify the optic nerve as normal or glaucomatous. The current process allows for misjudgments/errors in the interpretation of the CSLT image, inability to distinguish between actual and noisy images and variance in the diagnostic recommendations over a set of practitioners. The challenge, therefore, is to automate the analysis of CSLT images of the optic disc, in an objective and quantifiable manner, to support practitioners in the diagnosis and therapeutic management of glaucoma.

Researchers have analyzed optic nerve data and CSLT based images with varying results. Bowd et al [3], working with retinal tomography images applied forward and backward feature selection methods for training Multi Layer Perceptron (MLP), Support Vector Machine (SVM) and linear discriminant functions; Park et al [4] used correlation analysis and forward wrapper model to select features from optic disc data for training SVM classifiers; Swindale et al [5] used a wrapper model for feature selection to train SVM classifiers.

We have developed a data-driven Glaucoma Diagnostic Support (GDS) system that features the automatic interpretation of CSLT topography images of the optic nerve to support (a) the classification of the optic disc images to distinguish between healthy and diseased optic discs; (b) the identification of the sub-types of glaucomatous optic disc damage. This is to help further sub-classify the glaucoma patients in order to provide treatments in line with the specific morphological patterns of damage [6]; and (c) the visualization of the temporal progression of the disease for a patient over a period of time.

In this paper we present an automated approach to CSLT topography image analysis to support glaucoma diagnosis. Our multi-stage approach is a hybrid of image processing and data mining methods. In Stage 1, we apply image-processing techniques to CSLT images to derive image-defining features. In Stage 2, we apply data classification methods to the image's shape-defining features to develop classifiers that can discriminate between healthy and glaucomatous optic discs. An important task at this stage is feature selection whereby we select an optimal subset of image features that exhibit high image classification capabilities. In Stage 3, we apply data clustering techniques to the optimal subset of image-defining features in order to identify the different sub-types of glaucomatous images in the image data-set. The emergent image clusters are subsequently used to both visualize the progression of the disease and the identification of noisy optic nerve images. In Stage 4, we apply rule-induction techniques to the optimal subset of features to induce classification rules (not discussed here). These symbolic rules provide practitioners with a justification of the diagnostic recommendations by our image classifiers. For our experiments, we worked with 1257 tomography images taken at different time intervals from 136 subjects (51 healthy subjects and 85 glaucoma patients).

Methods

Figure 1 illustrates the functional design of our GDS system. We explain the methods developed for each processing stage.

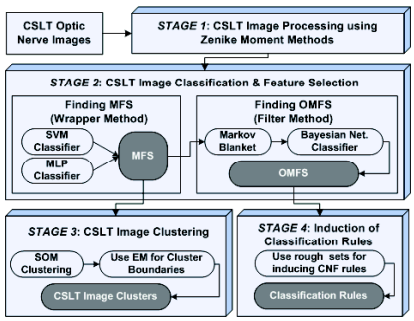


Figure 1 – Functional design of our GDS system

Stage 1: CSLT image processing

This stage involves the extraction of shape-defining features from CSLT images. These features are used to develop the image classification and clustering models. We use an image processing technique referred to as Moment Methods that describes the properties of connected regions in binary images as *Moment* features. We use Zernike moments [7] which use a set of complex polynomials to describe the image's properties by their order (n) and repetition (m) with respect to a digital image—the low order moments capture gross shape information and high order moments incrementally resolve high frequency information (representing detail) of the digital image. Two attractive features of Zernike moments for our purpose is that (a) moments can be made invariant to shifts, rotations and magnification changes; and (b) the optic nerve is centered in the image, thus avoiding the requirement for an independent segmentation stage in which the object is explicitly identified.

For each CSLT image we generated 254 Zernike moments, grouped in an incremental order ranging from 1 to 29—each group comprises a set of ordered moments. Low order moments capture fundamental geometric properties and high order moments represent detailed information of the image [7].

For efficient classification of CSLT images, it was important to select an optimal number of lower order moments. This is a non-trivial task because: (a) there is no objective measure to determine the exact number of (low order) moments needed to achieve high classification accuracy; and (b) there is no discernable relationship between the moments that can be utilized. Given these challenges, next we pursue feature subset selection in conjunction with image classification.

Stage 2: Classification of CSLT images

In the previous stage, we derived a 254 moment based representation for each CSLT image. In this stage, we pursue the classification of CSLT images based on a sub-set of low order moments. This stage therefore involves two tasks—i.e. firstly feature (sub-set) selection and secondly image classification. We have developed a two-pass image classification method that incorporates feature sub-set selection as an integral element (see Figure 1). In the first pass, MLP and SVM based wrapper models are simultaneously used to generate a *Moment Feature Subset* (MFS) consisting of low order moment features. In the second pass, we apply a Markov blanket filter method [8] based on an inferred Bayesian network to select the highly relevant moments from the MFS—i.e. the *Optimal Moment Feature Subset* (OMFS)—that offers reasonably high image classification despite using a small number of moments.

Pass I: Using MLP and SVM

In the absence of any guiding principle to determine the size of the MFS, we devised an accumulative feature subset selection strategy as follows: (a) Generate training-set by incrementally adding the next higher order moments to an existing training set. We exploited the intrinsic partitioning of the 254 moments in terms of their order ranging from 1 to 29. Therefore, feature subset1 included moments with *order2*, feature subset2 includes moments with order 2 and 3, and so on. In total 29 different training sets were generated, where each training set covered all the images based on the moment orders chosen to represent it; (b) Train both a MLP and a SVM classifier separately using the 29 training sets. In total, 29 different MLP and SVM classifiers were trained. For training the MLP and SVM, we partitioned the images so that 75% images were used for training and 25% images were used for testing. For training the SVM, based on the training data a 5-fold cross validation was performed to find the optimal hyper parameters: C and λ; and (c) Determine the classification accuracy of both classifiers, using the test images that are represented by the same number of moments as used to train the classifier.

The next step was to determine the size of the MFS and based on it to select the most efficient MLP and SVM classifier. Our objective was to select the largest possible number of moments without compromising the classification accuracy. To do so, we plotted the classification accuracy of both classifiers and then identified the highest accuracy point on the plot (i.e. with respect to n moment groups) just prior to a downward trend in the classification accuracy as a result of the inclusion of the next higher moment group. The most low order moment groups that achieved the highest classification accuracy were selected as the MFS. And, the MLP and SVM classifiers trained using the MFS were deemed as the most efficient.

A comparison of the classification accuracy trends for both the MLP and the SVM classifiers (see figure 2) shows that both classifiers exhibited a similar classification accuracy trend—i.e. they both start with a relatively high accuracy with the first moment group and then the accuracy drops with the addition of the next few moment groups. But later the accuracy starts to pick up again such that for the MLP it peaks when the feature subset constitutes the first 8 moment groups, whereas for the SVM the accuracy peaks for the first 11 moment groups. It is interesting to note that the classification accuracy with higher order moment groups is relatively low as compared to the peak achieved with just the lower order moments.

Based on classification accuracy trend for both classifiers (shown in figure 2), we determined the MFS to contain the first 11 moment groups—i.e. the first 47 moments. With 11 moment groups the SVM exhibited the highest accuracy and the MLP produced its second highest accuracy level.

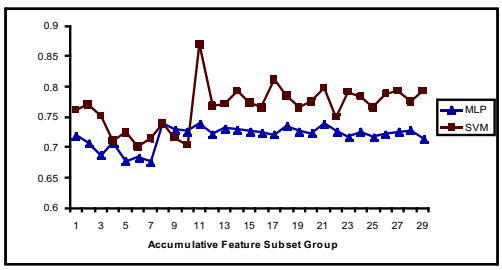


Figure 2 - Classification accuracy for both MLP and SVM

Pass II: Using Markov blanket and Bayesian network

In the second pass, we attempt to further reduce the size of the MFS in order to generate the OMFS that comprises only the highly salient moments. We use a filter model based on a Bayesian Network (BN) and the Markov blanket of the class label [8]. The choice of Markov blanket is guided by the observation that the correlation between most of the moments and their class label is weak, and the same is true for correlation between different moments. Hence, correlation based feature selection methods are not suitable here. We decided to use the Markov blanket approach as it considers every feature's probability dependence relationship during the learning procedure of the Bayesian network's structure.

In a BN where C_A is the set of children of node A, and Q_A is the set of parents of node A, the subset of nodes containing Q_A, C_A and the parents of C_A is called Markov blanket of A [8]. The Markov blanket of a specific feature is a subset of nodes in the BN; it comprises the feature's parent nodes, child nodes and all parent nodes of the child nodes. If we consider the class label node as the root node to learn a BN from data, then all nodes within the Markov blanket of the class node have probabilistic dependence relationship with it.

The steps to generate the Markov blanket were as follows: *Step 1*: We used the *K*2 algorithm to learn the BN. Initially, the 47 moments in the MFS were discretized using an entropy-based method, resulting in 29 moments to be discretized into a single value. These moments were removed from the MFS. Thus we were left with only 18 moments for training the BN. The following moments were retained: moments {1, 2, 5, 6, 7, 12, 16, 21, 23, 25, 27, 33, 36, 37, 43, 44, 45, 46}. *Step 2*: A BN was trained using the18 moments in their original order. Five-fold stratified cross validation was used to evaluate the classification accuracy (see table 2 for results). *Step 3*: The 18 moments were ordered based on the chi squared statistical test score x^2 between the moments and their class labels. The moments with the highest x^2 were: {1, 43, 16, 25, 21, 23, 6, 5, 36, 2, 27, 33, 37, 7, 46, 45, 44, 12}. A BN was learnt using the ordered moments (see table 2 for results). *Step 4*: From the BN learnt in step 3, we inferred the Markov blanket of the class label and found that only six (6) moments {1, 6, 16, 21, 37 46} were within the Markov blanket of the class label. These six moments were selected to form the OMFS. *Step 5*: In order to determine the classification capability of the selected OMFS we used them to train a BN. Next, 5-folds cross validation's classification accuracy was calculated (see table 2 for results) and it was noted that the OMFS offers quite high classification accuracy.

Stage 3: Clustering of CSLT images

In this stage we pursued the clustering of the CSLT optic nerve images, represented using the 47 moments in the MFS, to differentiate between the different subtypes of healthy and glaucomatous optic nerves. It may be noted that an important theme in glaucoma research is to develop an understanding of the large variation in the appearance of the optic nerve, both within groups of healthy subjects and in patients with glaucoma. It is therefore important, from a clinical standpoint, to recognise and differentiate between such patterns. However, the problem with the sub-classification of patterns of optic nerve damage is that it is a subjective task, giving rise to considerable levels of disagreement between trained experts. In this context, our aim was to develop an objective and automated method to characterize optic nerve images.

Our two phase clustering strategy was to: (a) partition the images into distinct clusters using Self Organizing Maps (SOM); and (b) draw clear and distinct boundaries around the clusters using the Expectation Maximization (EM) algorithm [9].

Phase A: Data clustering using SOM

We used a SOM to cluster the CSLT images based on the similarities between image shapes, where each cluster may represent a different subtype of healthy and glaucomatous optic nerves. The training of the SOM was conducted as follows: (i) we determined the topology of the SOM to be hexagonal lattice comprising 192 units that were arranged as 16 rows and 12 columns; (ii) The units were linearly initialized along the two greatest eigenvectors of the covariance matrix of the training data—i.e. images represented using the 47 moments in the MFS; (iii) The SOM was trained using a sequential training algorithm by first running a rough training phase comprising 100 epochs starting with a large neighbourhood radius of 12 that was linearly reduced to 3 with a learning rate of 0.5. This was followed by a second fine-tuning phase comprising 1000 epochs with a small initial neighbourhood radius 3 that was reduced to 1 with learning rate of 0.1. In both cases a Gaussian neighbourhood function was used and the learning rate function was set to be inversely proportional to the training epochs; (iv) Finally, we achieved a trained SOM that placed similar images into close proximity, thus leading to the image clusters. We applied principal component projection to the learnt SOM to determine its projection. This was followed by developing a U-matrix representation of the learnt SOM by spreading a colour map on the projection. Based on the visualization offered by the SOM, it was noted that the data was partitioned into discernable clusters.

Phase B: Defining the cluster boundaries

After determining broad clusters of CSLT images, in this phase we objectively determine the cluster boundaries. The processing was guided by our assumption that the dis-

tribution of the clusters within the learnt SOM is Gaussian. Therefore, we used the EM algorithm [9] as it is suitable to find distinct components in the case of Gaussian mixtures. Functionally, the EM algorithm initiates with an estimate of the number of components and their parameters. Our strategy was to maximize the likelihood of the optic nerve images into distinct clusters given the parameters and a maximum likelihood measure that indicated how well the Gaussian mixtures fit the data into clusters. We used a Bayesian Information Criterion (BIC) [9], where the best estimate (e.g., number of clusters) was chosen based on the highest BIC value.

To achieve the cluster boundaries, using the EM method with BIC, we initialized the EM using 10 random re-starts method, and then selected a parameter setting to maximize the log-likelihood of our clusters from the SOM. EM clustering was performed for different number of clusters. Table 1 shows that the maximum BIC is achieved when K = 4. Hence, we determined that given the learnt SOM there are 4 clusters—one cluster represents health images and the three clusters for sub-types of glaucomatous images—in it that best fit the data (see Table 1). To finalize the cluster boundaries for the 4 clusters, we calculated the assignment probabilities of each CSLT image to all the cluster labels, the cluster label with the highest probability value was assigned to the CSLT image. Figure 2a shows the SOM with the emergent clusters, the clusters are coded using grey scale for visualization purposes.

Table 1 – Number of clusters vs. BIC values

K	2	3	4	5	6	7	8
BIC	29100	30409	31354	30516	29125	27456	25486

Evaluation and discussion

In this section we present the evaluation results for the various methods developed for stages 2 and 3 of our GDS system.

Evaluation 1: Evaluating CSLT image classification

Table 2 presents the CSCLT image classification accuracy for the different classifiers trained in phase 2, using test images not previously seen by the classifiers. It is interesting to note that both the MLP and the SVM classifiers offered higher accuracy with the MFS as compared to the original 254 moments. This vindicates our hybrid feature sub-set selection strategy, and also confirmed the theoretical assumption that low order moments contain more shape information that is relevant for classification as compared to the information content of high order moments. In the second pass, we determined that the MFS could be further reduced to just 6 moments—i.e. the OMFS—without compromising the classification accuracy. The highest accuracy for MFS was offered by the SVM—i.e. 86.96%. The highest accuracy for the OMFS was 83.82% offered by a BN. Therefore, the compromise in classification accuracy is just 3 %, yet the gain in computational efficacy is quite significant. Note that the BN (with Markov Blanket) offers the most optimal classification results when compared with both MLP and SVM trained on the OMFS. We therefore selected the BN classifier trained with the OMFS to distinguish between healthy and glaucomatous optic nerves.

Table 2 - Classification accuracy for different classifiers

Feature Subset Size	**Classifier**	**Accuracy**
Pass I		
254 moments	MLP	72.88%
254 moments	SVM	77.50%
47 moments in MFS	MLP	74.00%
47 moments in MFS	SVM	86.96%
Pass II		
18 moments (original order)	BN	77.21%
18 moments (chi^2 order)	BN	80.88%
6 moments in OMFS	**BN**	**83.82%**
6 moments in OMFS	SVM	80.26%
6 moments in OMFS	MLP	72.84%

Evaluation 2: Examining the CSLT image clusters

Evaluation of the clustering stage involved mapping a series of optic nerve images for individual patients (i.e. test cases with explanations provided by experts) onto the SOM and noting the Compactness Factor (CF) between the activated units. The CF measures how close the images are with respect to each other in terms of the average distance between the centroid of all active units. The CF is an objective measure for evaluating the clustering goodness based on our initial observation that for a patient the series of optic nerve images are quite similar over a period of time; over time the differences are quite minute and should not lead to large variations between consecutive images. This implies that when visualizing the optic nerve images for a subject, the active units should be in close proximity and therefore yield a low CF.

Figure 2a show that the results for patient 4209643, and it maybe noted that the 7 optic nerve images, taken over a period of time, map on to a single SOM unit resulting in a compactness factor of to 0. The numeral within the active unit shows the number of images mapping on to that unit. This demonstrates the best possible clustering outcome as the learnt SOM recognizes the similarity between all the 'healthy' optic nerve images for this patient. Figures 2b shows the 11 optic nerve images of patient 112455 being mapping on to 4 neighboring SOM units within one cluster, with a compactness factor of 0.20808. This result again implies the close proximity of the images for this patient. These results are in line with the visual observations of these images by experts, who also concurred that the images for these patients are quite similar in shape.

We use the learnt SOM to visualize the disease progression for a patient over a period of time. Images taken over time for a patient were mapped onto the SOM. The pattern of the active units indicated the potential progression of the disease from one cluster to another, where each cluster may represent images of a specific glaucoma sub-type. In Figure 3a the images fall into two adjoining clusters, and the path across the clusters suggests the progression of the

disease from one sub-type to another. Figure 3b shows the progression over time.

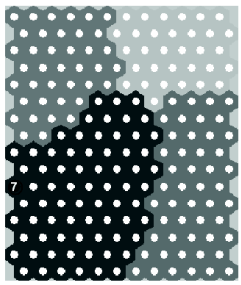


Figure 2a – SOM showing all images mapped to a unit

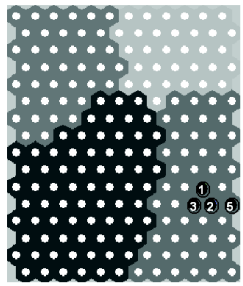


Figure 2b – SOM showing all images mapped in one cluster

Evaluation 3: Visualizing disease progression over time

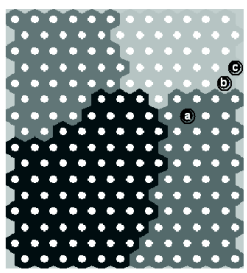


Figure 3a - SOM showing the dispersion of images over two adjoining clusters

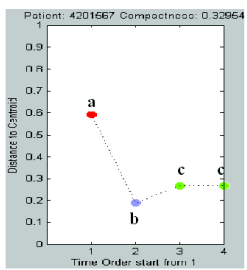


Figure 3b – The disease progression path. Note the high CF between the images

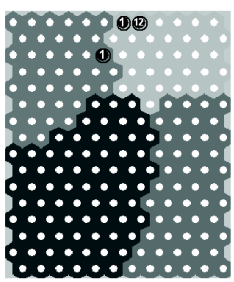


Figure 4a - SOM showing a noisy image that is distant from the other images.

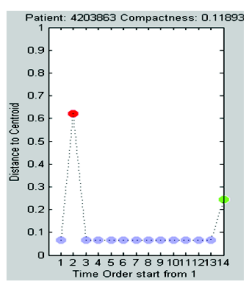


Figure 4b – The spike shows that the 2nd image is noisy, as it does not follow the pattern.

Evaluation 4: Identifying noisy CSLT images

We used the learnt SOM to identify noisy CSLT images that typically occur due to various factors related to the capture of the optic nerve image. With the knowledge that consecutive images do not manifest drastic changes, if an image is noted to be significantly dissimilar from its neighbors it can be regarded as a noisy image. At present there are no objective means to identify noisy CSLT images. Figure 4 (a-b) shows 14 images for a patient, where the 2nd image is identified as a single noisy image because it is in a different cluster, whereas the remaining images all map onto just two other units that are very close to each other.

Concluding remarks

We presented a data mining framework to objectively analyze medical images, and applied it to investigate glaucoma. The novel features of our approach are that: (a) we process CSLT images to derive shape information by using image processing techniques. This is in contrast to the traditional approach of using morphological features to analyze CSLT images; (b) we have developed a feature selection strategy that identifies the most salient image-defining features without compromising the classification accuracy; and (c) we are able to visualize the CSLT images in terms of clusters of similar images. These clusters provide an opportunity to visualize the dispersion of multiple observations for a subject, and we show how this information can help to (i) determine a potential progression of the disease due to changes in the optic disc over time; and (ii) identify noisy CSLT images. We believe that our framework takes a step towards the automated and objective analysis of optic nerve images to support glaucoma diagnostics.

Acknowledgments

This project is supported by a research grant from Nova Scotia Health Research Foundation, Canada.

References

[1] Hoskins H, Kass M. Primary open-angle glaucoma, Becker-Shaffer's Diagnosis and Therapy of the Glaucomas. The C.V. Mosby Company, St.Louis, 1989: 277-307.

[2] Zinser G, Wijnaendts-van-Resand RW, Dreher AW. Confocal laser tomographic scanning of the eye. Proc. SPIE 1161, 1980: 337–344,

[3] Bowd C, Cban K, Zangwill LM, Goldbaum MH, Lee T, Sejnowski TJ. Comparing neural networks and linear discriminant functions for glaucoma detection using confocal scanning laser opthalmoscopy of the optic disc. Investigative Opthalmology & Visual Science, Vol. 43 (11), 2002.

[4] Park J, Reed J, Zhou Q. Active feature selection in optic disc nerve data using support vector machine. IEEE World Congress on Computational Intelligence, 2002.

[5] Swindale NV, Stjepanovic G, Cbin A, Mikelberg F. Automated analysis of normal and glaucomatous optic nerve head topography images. Investigative Opthalmology & Visual Science, Vol. 41 (7), 2000.

[6] Broadway DC, Nicolela MT, Drance SM. Optic disc morphology on presentation of chronic glaucoma. *Eye* 17 (6), 2003: 798-799.

[7] The CH, Chin RT. On Image Analysis by the methods of moments. IEEE Trans. Pattern Analysis by Machine Intelligence 10(4), July 1998: 96-513.

[8] Hruschka Jr. ER, Hruschka ER, Ebecken NEF. Feature selection by Bayesian network. Canadian AI Conference, LNAI 3060, Springer Verlag, 2004:370-279.

[9] McLachlan G, Krishnan T. The EM algorithm and extensions. New York: John Wiley & Sons, 1997.

Address for correspondence

Syed Sibte Raza Abidi,
Faculty of Computer Science,
Dalhousie University,
Halifax, B3H 1W5, Canada.
Email: sraza@cs.dal.ca

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Intelligent Querying and Exploration of Multiple Time-Oriented Medical Records

Denis Klimov[a], Yuval Shahar[a]

[a] *Medical Informatics Research Center, Ben Gurion University of the Negev, Israel*

Abstract

Querying and analyzing multiple time-oriented patient data is a key task during medical research, clinical trials or the assessment of the quality of therapy. In this paper, we present several aspects of the VISITORS system, which includes knowledge-based tools for graphical querying and exploration of multiple longitudinal patient records. We focus on the syntax and semantics of the knowledge-based aggregation query language for multiple time-oriented patient records, and on the graphical query-construction interface. The query language assumes an underlying computational method for deriving meaningful abstractions from single and multiple patient records, such as we had previously developed. The aggregation query language enables population querying using an expressive set of constraints. By using our underlying temporal mediator architecture, the time needed to answer typical temporal-abstraction aggregation queries on databases of 1000 to 10000 patients was reasonable.

Keywords:

intelligent visualization, temporal abstraction, multiple patients, medical informatics, Human-Computer Interfaces

Introduction: Knowledge-based exploration of multiple time-oriented records

A key task facing clinicians and medical researchers is the analysis of time-stamped, longitudinal medical records, in particular records of multiple patients. This capability is necessary to support, for example, quality assessment tasks, analysis of clinical trials, or the discovery of new clinical knowledge. Although the task of assessing patient data has been mostly solved through the increasing use of Electronic Medical Record (EMR) systems, there still remains the task of intelligent processing of multiple time-oriented patient records, including the capability for interactive exploration of the results. Standard means, such as tables, temporal statistical tools, or more advanced temporal data mining techniques, are often insufficient or can help only in particular cases.

To solve the computational aspect of this problem, we have been using the *knowledge-based temporal abstraction* (KBTA) method [1] for automated derivation of meaningful context-specific interpretations and conclusions, called *temporal abstractions*, from raw time-oriented patient data, using a domain-specific knowledge-base (KB). In general, the KBTA method is defined as follows: The input includes a set of time-stamped parameters (e.g., platelet, red blood-cell (RBC), and white blood-cell (WBC) counts) and events (e.g., bone-marrow transplantation (BMT)), which create the necessary interpretation context (e.g., the therapy protocol used). The output includes a set of interval-based, context-specific parameters at the same or a higher level of abstraction and their respective values (e.g., a period of nearly 3 months of grade 0 bone-marrow toxicity in the context of that therapy protocol).

Furthermore, the output temporal abstractions can be efficiently visualized. The KNAVE-II system, which we developed previously [2], supports the visualization and exploration of raw data and derived temporal abstractions for an individual patient or small number of patients. Evaluation of the KNAVE-II system in the oncology domain [3] has demonstrated that, by using KNAVE-II and its underlying temporal abstraction computational architecture, physicians can more quickly and more accurately answer clinical queries about patients.

However, to analyze clinical trials, or for quality assessment purposes, an aggregated view of a group of patients is more effective than exploration of each individual record separately. In addition, certain patterns can only be discovered through the analysis of multiple patients. Therefore, we have extended the KNAVE-II system into a system called VISualizatIon of Time-Oriented RecordS (VISITORS) [4] which supports the visualization of a group of time-oriented records at different levels of abstraction.

The following three important features distinguish the VISITORS framework from other exploring data tools :

1. Time-oriented data are graphically displayed and explored for both individual and multiple patients.
2. The temporal dimension is a first class citizen. It can be explored in various granularities, such as hour, day, and month. We also support a calendar timeline and a time-line relative to special events (e.g., the six months following a particular intervention).
3. The computational reasoning supports not only a view of raw time-oriented data and their statistics but also a meaningful summarization of the raw data, based on the temporal-abstraction domain ontology and the KBTA computational mechanisms. The exploration interface is also based on that ontology, which supports a semantic exploration of the data and enables navigation of semantically related raw and abstract concepts. For example, the user can explore graphically an instance of a pattern that was derived by the KBTA pattern-detection inference mechanism and view all the

abstract components and raw data from which the pattern was derived.

In this paper, we explain in detail the syntax and semantics of the VISITORS query model, which directly affect the semantics of the computational and display modules. We also introduce several of the graphic modules which we have implemented to assist users in the interactive definition of temporal aggregated queries and exploration of multiple time-oriented records. Finally, we provide results of a preliminary functional evaluation and discuss the implications of the VISITORS framework.

Methods

General architecture

The VISITORS system is an intelligent interface to a distributed architecture specific to the tasks of querying, knowledge-based visualization, and interactive exploration of time-oriented data. We assume that the necessary elements of the temporal abstraction framework (shown in Figure 1 by striped lines) are available. Figure 1 describes the overall architecture: End users (clinicians) interact with *Query Builder* of VISITORS to submit time-oriented queries regarding patients. The Temporal Abstraction Mediator, for example our previously described systems, the goal-directed IDAN Mediator [5], integrates the relevant data and knowledge from the appropriate sources, indicated by the user, to answer queries regarding raw data or to derive using a Temporal Abstraction Service the abstract time-oriented concepts. The resultant data can be visualized and explored by the user.

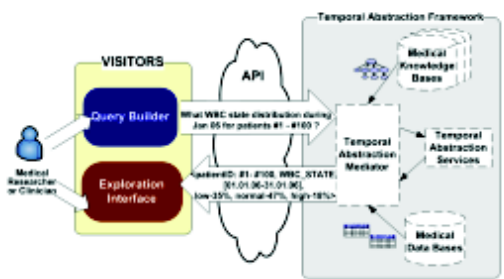


Figure 1- The distributed VISITORS architecture

Semantics of the temporal aggregation queries

We assume that the clinical time-oriented databases (DBs) have the following virtual patient record structure: the patient identification data (e.g., patient ID), name of entity (e.g., measured parameter, medications, interventions, etc.), time of laboratory test or medication and value. For example:

<John Smith,WBC,11.01.06 13:45:00,6.7*103 cells/ml>,

In our work we have designed and developed a formal query language, which is ontology-based, i.e., we use domain knowledge to formulate and display the queries. We distinguish three types of aggregated queries.

Select patients query

This query retrieves a list of patients from selected database who satisfy a set of constraints, defined by user:

GetPatients (KB, DB, <PatientsConstraints>) $\Rightarrow$ *<patients>*,

where *GetPatients* is an external procedure that queries the selected *DB*, *KB* is an appropriate knowledge base that includes the parameters definition used in *<PatientConstraints>* - a set of complex conditions defining criteria for patient selection, and *<patients>* is a resultant list of identification data of patients who satisfy the set of *<PatientConstraints>*.

The *<PatientConstraints>* aspect of the query includes the list of Boolean and temporal conditions of three types:

- **Demographic** or **non-temporal constraints** (i.e., only the last value is relevant) such as patient's ID, age, sex, physician, etc. The user can define Boolean constraints among the attribute's values (i.e., OR/AND logical operators). However, we do not recommend using complex Boolean expressions, since they make understanding why a patient was included in the output less intuitive. The logical operator NOT was omitted for a similar reason.
- **Time and value constraints** for both raw and derived concepts. To construct the query, the user can define constraints on the value of a concept, its duration, and its start/end points (See Figure 2). Both absolute (i.e., calendaric) and relative (i.e., measured from a reference event) timelines are supported. Defining pair-wise temporal relations between interrelated concepts is supported using Allen's temporal logic relation as well as relations among specific time points (i.e., start/end).

  We use the conjunction logical operator between the time and value constraints. Note that the computational functions and procedures that enabled us to derive the abstract parameters are defined in the knowledge base as part of the domain ontology. The name of the KB is specified in the query.

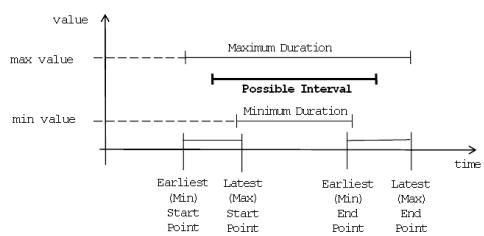


Figure 2-Time and value constraints for concept

- **Statistical Constraints** enable the user to aggregate and filter the patient's data on the basis of a specific statistical function. Using such constraints, the user is able to investigate who are the patients in the database who have specific values (or a range of values) within a given statistical range of threshold values. For example: "Select all patients whose state of WBC was derived as "low" or "very low" during more than 25% of the period Jun 1 – Oct 31 2006".

Figures 3 & 4 show an example of constructing a *Select Patients Query* with both demographic and knowledge-based constraints, whose informal definition is "Select all male patients, either younger than 20 or older than 70, whose hemoglobin (HGB) state was abstracted at least as

"moderate low" or higher, during at least seven days, starting at least two weeks after the allogenic bone marrow transplantation(BMT),and whose WBC counts were increasing during the same period." That is, either young or old patients whose bone marrows have been recovering following the BMT procedure.

The bottom part of the interfaces shown in Figures 3 and 4 displays the query that is automatically and incrementally being created from the user's graphical specification (displayed in the top part). The graphical interface used for query construction has a similar structure to the main exploration interface of the VISITORS system, which we discuss in the next section (e.g., the ontology browser is displayed on the left side; the panel display used for query definition is similar to the one used for data display, etc.). The highlighted rectangular area denotes the ranges of the time and value constraints on the patient data or abstractions relevant to the query.

Select time intervals query

Given a set of time-oriented patient data, this query returns a list of time intervals that satisfy the constraints defined by the user. In other words, the goal of this query is to find *when* a certain portion of the patients has a specific value or value in a predefined value range.

Formally, the Select Time Interval Query has the following structure:

```
GetTemporalIntervals (KB,DB,
<Constraints>, [refPoint]) ?
     relation*(<start_time, end_time>),
```

where *GetTemporalIntervals* is an external procedure that queries the selected *DB* and *KB* is an appropriate knowledge base that includes the parameters definition used in *<Constraints>*. The output *relation*(<start_time, end_time>)* is a data structure that includes a list of temporal intervals distinguished by *start_time* and *end_time* time points. The optional parameter *refPoint* requires that the output be calculated from the denoted reference point (e.g., a key clinical procedure).

<Constraints> is a logical expression without nested parentheses of the following atomic constraints related by conjunction logical operator:

```
<Concept,    min_thresh,    max_thresh,
min_value, max_value>,
```

where *Concept* is a concept name, the *min_thresh* and *max_thresh* are minimal and maximum thresholds values in percents of patients, and *min_value* and *max_value* is a range of values of the selected *Concept*.

Figure 3-Definition of a query's demographic constraints

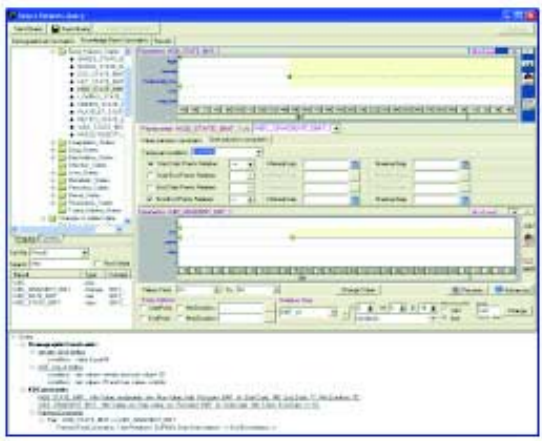


Figure 4-Definition of a query's knowledge-based constraints

For example, a typical *Select Time Interval Query* is: "Find [relative] time intervals following the BMT, during which the WBC count was increasing, and the state of the Platelet count was "normal" or higher, for more than 30% of the patients." Its formal XML based definition would be:

```
<SelectTimeIntervalsQuery>

  <Concept name="Platelet state"
min_value="normal"

           max_value="very high"

           min_theshold="30"
max_theshold="100"/>

  <Concept name="WBC
gradient"min_value="increasing"

           max_value="increasing"

           min_theshold="30"
max_theshold="100"/>

  <ReferencePoint name="BMT"
event="last"/>

</SelectTimeIntervalsQuery>
```

Get patients data query

Given a list of patients ID's and, optionally, a list of time intervals, the query retrieves the patients' raw data, or the derived temporal abstractions of one selected concept within the selected time intervals for the selected patients. The default patient list is all of the patients in the DB, and by default there are no time-interval constraints. The formal expression of the query is the following:

```
GetDataConcept(KB,DB,Concept,<patients>,[<t
ime intervals>])?
       relation*(<patient_n,
start_time_n,m, end_time_n,m,value_n,m>),
             1  n  N, 1  m  M,
```

where *GetDataConcept* is an external procedure that queries the selected DB, *KB* is an appropriate knowledge base that includes the *Concept* definition, *<patients>* is a list of patients identification data, and *<time intervals>* is an optional parameter that constrains the time of the returned

data. The output *relation*(<patient$_n$, start_time$_{n,m}$, end_time$_{n,m}$, value$_{n,m}$>)* is a data structure that includes the time-oriented data of *N* patients in <*patients*> list with *M* data records for each patient.

In contrast to the previous two queries, the *Get Patients Data Query* does not have a separate interface. This query is constructed automatically during the users' exploration of the patient's data in the main interface of VISITORS.

The main interface in VISITORS

We described previously several visualization tasks in VISITORS [4]; here we intend to explain only the main design principles of the interface.

The main interface in the VISITORS system is divided into three logical parts (See Figure 5):

The top panel (A) is used for the patients selection tasks. The user can select previously retrieved groups from the table, choose patients from a DB, input the patient ID by himself, or construct a new *Select Patients Query*.

The middle small panel (B) is used for time interval selection. The user can use the previously returned intervals, define a new interval, or construct a new *Select Time Intervals Query*. Note that both calendaric and relative timelines are supported. The main part of the interface (C) is used to explore the patients' time-oriented data. The left side includes a browser to the clinical domain ontology, retrieved on the fly from the relevant domain KB. Clicking on a concept node in the ontology tree displays the data of that concept for the selected group of patients. In this case, the user explores the data of a group of patients named *My patients* (58 patients), previously retrieved by the *Select Patients* query. The 1st panel from the top displays all of the WBC laboratory test values during March 95. The top (red) line represents the daily maximal value of the WBC count. The 2nd panel shows the daily mean value of HGB for each patient during 1995. The top (red) line and bottom (blue) line represent, respectively, the monthly maximal and minimal patient HGB values within the group. The 3rd panel displays the distribution of the WBC state-abstraction values for each month of 1995. For example, in Aug 95, 29 % of patients in the group have had a *moderately_low* value. The bottom panel displays statistical and temporal associations within the specific time period among the selected raw-data or abstracted concepts. In this case, WBC and RBC values are displayed over March 1995, and the Platelet and HGB values over 1995. Only 25 patients in the group have data in the selected time intervals. Values of all parameters for each patient are connected by lines.

The functional evaluation

We are currently performing a functional and usability evaluation of the *Query Builder* module in an oncology domain. Evaluators were asked to construct eight *Select Patients* and two *Select Time Intervals* queries at different difficulty levels (Examples of queries are showed in Table 1).

Figure 5-The main VISITORS interface

Table 1 – Examples of aggregated queries

Complexity	Examples of Queries
Select Patients Query	
Easy	Find male patients older 50 who have had during Sep 95 WBC laboratory test value less than *4000 cells/ml*.
Moderate	Find all patients who had the HGB state value in the value range "*low*" to "*normal,*", within the following time constraints: the episode of HGB starts between 3 to 4 days after allogenic BMT and its duration is at least 20 hours but no more than 60 hours.
Hard	Select male patients, whose bone-marrow toxicity grades have been decreasing for at least seven days during the period starting from 2 weeks after the allogenic BMT procedure, and whose liver toxicity grades were also decreasing during the same period.
Select Time Intervals Query	
Easy	Find time intervals when the state of WBC was considered as "normal" or higher for less than 30% of the patients
Hard	Find time intervals following the BMT procedure during which more than 20% of the patients had "*moderate*" anemia, and more than 30% of the patients had "*low*" or "*very low*" WBC counts

In all tests we used a retrospective DB of more than 1000 oncology patients after bone-marrow transplantations (BMT), who were followed for two to four years. In addition, we assessed the performance of the overall architecture when answering these queries. Results show that performance times using IDAN architecture were reasonable: several seconds for answering aggregated queries with demographic and time and value constraints for raw and abstracted parameters. We intend also to evaluate the system in the diabetic domain.

Discussion

In this paper we have presented the novel ontology-based multiple records query language, which enables users to construct a new query and retrieve a set of relevant patients or time intervals, using a broad set of constraints. Such aggregated queries, and the graphical *Query Builder* tool used to construct it, are part of the VISITORS system that enables clinicians to query, visualize and explore both raw time-oriented medical data and meaningful interpretations (including complex temporal patterns), derived from the these data, based on a domain knowledge base. The main advantage of our system is successful integration of different methods, such as information visualization and knowledge-based temporal reasoning.

In previous studies the time-oriented aspect of querying was commonly addressed by adding a temporally-oriented extension to standard SQL tools [6, 7]. However, these systems are limited in the temporal analysis by SQL capabilities. Moreover, they do not include the temporal reasoning mechanisms. In the area of visual querying, several tools have been proposed. The TimeFinder system [8] is a visual exploration and query system for exploring time-series data sets, based on a direct manipulation metaphor. Chittaro and Combi [9] provide visual representations framework of temporal intervals and relations. However, these techniques focus on query and exploration only of raw longitudinal data. Attempts to support intelligent query and retrieval are provided in the TimeLine system [10], and in knowledge-based spatial temporal query language (KSTL) [11] which support, however, only the medical imaging domain. Thus, we aimed to develop a domain-independent visual query and exploration system that enables clinicians to explore multiple longitudinal patient records, in an intelligent manner that supplies querying both raw data and their *knowledge-based meaningful interpretations*.

Acknowledgments

This research was supported by Deutsche Telekom Company and Israel Ministry of Defense award no.89357628-01

References

[1] Shahar Y. A framework for knowledge-based temporal abstraction. *Artificial Intelligence*, 90(1-2), 1997.

[2] Shahar Y, Goren-Bar D, Boaz D, Tahan G. Distributed, intelligent, interactive visualization and exploration of time-oriented clinical data and their abstractions. *Artificial Intelligence in Medicine* 38 (2), 2006, pp. 115-35.

[3] Martins SB, Shahar Y. Evaluation of KNAVE-II: A tool for intelligent query and exploration of patient data. *Proc. of Medinfo 2004*.

[4] Klimov D, Shahar Y.A framework for intelligent visualization of multiple time-oriented medical records. *AMIA Annu Symp Proc. 2005*, pp. 405-9.

[5] Boaz D, Shahar Y. A distributed temporal-abstraction mediation architecture for medical databases. *Artificial Intelligence in Medicine* 34 (1), 2005, pp. 3-24.

[6] Nigrin D, Kohane IS. Temporal expressiveness in querying a time-stamp–based clinical database. *J Am Med Informatics Assoc*, vol. 7, 2000, pp.152-63.

[7] Combi C, Cucchi G. GCH-OSQL: a temporally-oriented object-oriented query language based on a three-valued logic. *Proc of the 4th International Workshop on Temporal Representation and Reasoning*, 1997

[8] Hochheiser H, Shneiderman B. Visual Specification of Queries for Finding Patterns in Time-Series Data. *Proc. of Discovery Science 2001*, 2001, pp. 441-46.

[9] Chittaro L, Combi C. Visualizing queries on databases of temporal histories: new metaphors and their evaluation. *Data and Knowledge Engineering*, 44, 2003, pp. 239-264.

[10] Aoyama D, et al. TimeLine and visualization of multiple-data sets and the visualization querying challenge. *Journal of Visual Languages & Computing* 18 (1), 2007, pp.1-21.

[11] Chu W, Alfonso F, and Ricky K. Knowledge-based image retrieval with spatial and temporal constructs. *IEEE Transactions on Knowledge and Data Engineering* 10(6).

Address for correspondence

Mr. Denis Klimov, Tel: 972-8-6477160, email: klimov@bgu.ac.il
Department of Information Systems Engineering, Ben-Gurion University of the Negev, P.O.Box 653, Beer Sheva, 84105, Israel.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Analyzing Web Log Files of the Health On the Net HONmedia Search Engine to Define Typical Image Search Tasks for Image Retrieval Evaluation

Henning Müller [a], Célia Boyer [b], Arnaud Gaudinat [b], William Hersh [c], Antoine Geissbuhler [a]

[a] *Medical Informatics Service, University and Hospitals of Geneva, Geneva, Switzerland*
[b] *Health On the Net (HON), Geneva, Switzerland*
[c] *Oregon Health and Science University (OHSU), Portland, OR, USA*

Abstract

Medical institutions produce ever-increasing amount of diverse information. The digital form makes these data available for the use on more than a single patient. Images are no exception to this. However, less is known about how medical professionals search for visual medical information and how they want to use it outside of the context of a single patient. This article analyzes ten months of usage log files of the Health on the Net (HON) medical media search engine. Key words were extracted from all queries and the most frequent terms and subjects were identified. The dataset required much pre-treatment. Problems included national character sets, spelling errors and the use of terms in several languages.

The results show that media search, particularly for images, was frequently used. The most common queries were for general concepts (e.g., heart, lung). To define realistic information needs for the ImageCLEFmed challenge evaluation (Cross Language Evaluation Forum medical image retrieval), we used frequent queries that were still specific enough to at least cover two of the three axes on modality, anatomic region, and pathology. Several research groups evaluated their image retrieval algorithms based on these defined topics.

Keywords:

log files analysis, image retrieval evaluation.

Introduction

An increasing amount of medical information is being produced digitally, making it available for further processing and use, i.e., for teaching and research. Much of the produced data and experiences from past cases can be used to create tools for diagnostic decision aid. A great deal of medical information is also available on the Internet, as there are increasing requests for medical information by patients and professionals [1]. MedLinePlus is one example of a repository created to inform non-professionals, patients searching for information. Another example is Health On the Net (HON), which develops quality criteria for medical web pages and has an accreditation service for pages adhering to several quality criteria. HON also runs web search engines for medical web content aimed at patients and medical professionals with a multilingual search interface[1] [2]. Much research has been done on the searching of medical texts [3] but less on how images are used and searched for, although the amount of image data being produced is rising [4]. Many medical image databases are available within institutions, mainly for teaching, but some are also made available on the Internet. These include Casimage, HEAL (Health Education Assets Library), MedPix, and the Pathopic datasets. MIRC[2] (Medical Image Resource Center) is an initiative of the Radiological Society of North America (RSNA) to unite teaching files under a single interface. These databases contain thousands of annotated images. Unfortunately, the images are only rarely indexed in search engines such as Google as they are usually only available through the search in database fields. Another problem is that the annotation is often incomplete and information on the image modality is not always given. A search for "lung CT" with Google image search in October 2005 brought 160 results, about half of them lung CTs. The abovementioned databases alone contain several thousand lung CTs.

Outside of medicine, visual information retrieval has been an extremely active research domain for more than 15 years [5]. Studies on domain-specific user requirements have been performed, for example for journalists searching for images [6] or in the cultural heritage domain [7]. In the medical field, visual information retrieval has been proposed many times as extremely useful [8, 9]. Still, most research has a limited focus on retrieval for one particular group of images [9]. Although this might be a domain with high potential impact, teaching and research are more likely to profit first from possibilities to browse very large and diverse image collections by visual properties. In the context of ImageCLEFmed [10], a challenge evaluation for medical image retrieval, two surveys were performed among medical image users [11, 12] to find out more about typical information needs and search tasks. CLEF (Cross-Language Evaluation Forum) is a challenge evaluation for retrieval of multilingual information. ImageCLEFmed in particular evaluates the quality of retrieval from multilingual medical image retrieval available on the Internet. The

1 http://www.wrapin.org/ & http://www.hon.ch/HONselect/
2 http://mirc.rsna.org/

Selected for best paper award.

surveys include five user groups: medical professionals for diagnosis, teaching, and research as well as medical students and librarians. The goal of the work descried in this paper was to create realistic search tasks for ImageCLEFmed[3] based on information needs of web users. The analysis resulted in 30 search tasks used by participating research groups. Among the techniques used was analysis of log files, an active research domain [13], mainly to analyze web page design.

Materials and methods

Used data sets

The data used for this study were log files containing query terms of the HONmedia[4] search engine. The examined period of queries included ten months, from January 1, 2005 to October 31, 2005. This period was sufficient for a representative evaluation of search terms. Variations of search frequency or quality over the months were not part of our analysis. The original data set contained 53'970 queries. With each automatically extracted query term, the date and time of the query was stored. It was also stored whether the query was directly done via the HONmedia interface or referred to from Google towards HONmedia search. Many queries were in French, as the French-speaking medical community frequently uses the HON query engine. It was not possible to perform an automatic translation of the topics, as language detection is hard with only very few words. Other languages identified for the queries were English, German, Spanish, and Italian.

Pre-treatment of the data and evaluation techniques

The analysis of the data was done on a Linux computer using Perl to analyze the text files. The original data sets were transferred to pure text and the information on time and date of the query were discarded. Perl was used mainly to pre-treat the data. As data were extracted automatically and as robots perform queries on web interfaces there are many different formats for queries (sometimes broken), plus a variety of international character containing umlauts and accented characters sets that need to be combined.

Results

The data contained two groups of queries, queries directly asked via HON and queries forwarded via Google. These groups were treated separately. A total of 37'293 queries were directly performed via HONmedia and 16'677 were forwarded via Google.

Text normalization

First, normalization was necessary for the text to remove differences in coding of the strings, parameter options transmitted and for broken queries containing graphical symbols. We did not treat the word order in the queries. The steps were mainly based on a manual analysis of the data:

- Unify coding issues, to remove accents, Umlauts, national symbols, and any sort of non-text: –"()+–.
- Remove commands and options send by web robots or search engines.
- Remove URLs or fragments of URLs.
- Convert all characters to lower case.
- Change plural of frequent terms such as "images".
- Remove frequent terms to define the target media: image(s) (5'796), media (512), video(s) (334).

Over 100 rules for normalization and removal were defined and applied to clean the data. Even after the removal steps, it was apparent that an extremely large number of different queries remained. In total, there remained 5'365 different unique queries (of 16'677) for the Google queries and 17'643 different HON queries (of 37'293). This meant that almost half the queries were unique being asked only once, which made a systematic evaluation of the entire dataset hard. The number of words per query was small. Google queries contained an average of 2.01 words in our study and HON queries 1.50 words, after removing the words image, video and media. This resulted in 191 empty queries for Google and 150 for HON. The same number of queries contained only a single character.

Removal of unclear queries

After term normalization, it became clear that there are queries unimportant for further analysis. First, a group of queries concerned sexually explicit queries: In the Google queries, the following frequent terms were removed: xxx (334 times). For HON the following terms were removed: penis (114), vagina (108), breast (102), sex (65), clitoris (32), gynecology (24). Another group of queries implicitly contained similar ideas; for Google these were: accouchement (childbirth, 143), cesarienne (33). For HON: home childbirth (239), nurse (130), birth (69). Third, another group of queries were processed to remove those not containing a precise information need, some of them, such as the term "search," were simply placed by web robots trying to access information stored in web-accessible databases. For Google this included the following terms: medical images (508 times), HON (116), health (62), medical illustrations (32), repository (30). For HON, these terms included: search (1493), medical images (79), doctor (70), anatomy (65).

Most frequent queries and terms

After normalization and removal of queries, we analyzed the most frequent remaining terms. Table 1 shows the most frequent remaining terms forwarded from Google. This list contains very specific medical search requests, from specialists rather than patients. Most of the terms are in French, actually all of the most frequent 20. The specialized nature of the terms and the fact that they are in French can be explained with the fact that only these technical queries link towards HONmedia.

3 http://ir.ohsu.edu/image/

4 http://www.hon.ch/cgi-bin/HONmedia/

Table 1 - Most frequent terms forwarded from Google

Term	*Frequency*
Nerf sciatique	154
Kyste pilonidal	76
Leucemie aigue myeloblastique	72
Glossite exfoliatrice marginee	67
Fracture humerus	66
Grenouillette sublinguale	60
Hematome sous dural	57
Polype nez	56
Appendice xiphoide	53
Leucomalacie periventriculaire	51
Leucemie	46
Purpura rhumatoide	46
Scarlatine	44
Hematome retroplacentaire	40
Kyste thyreoglosse	39
Leucemie myelomonocytaire chronique	39
Leucoplasie	38
Apophyse odontoide	37
Hidradenite	37
Scoliose	34

Table 2 shows the most frequent terms directly queried with HONmedia. These terms are more likely to be from patients than specialists. The first 20 contain only a single word. More terms are in English than in French, actually all top 20, whereas a large number of the less frequent terms are in French. Most terms are of two groups: Terms describing an anatomic region or a disease. Only other terms found in the most frequent 20 are concerning symptoms or a treatment in the largest sense, such as *injection, bacteria and pain.*

Table 2 - Most frequent terms from the HONmedia search

Term	*Frequency*
Heart	381
Asthma	242
Brain	211
Diabetes	160
Liver	101
Cancer	98
Marfan	93
Kidney	77
Lung	69
Knee	69
Injection	67
Bacteria	64
Eye	60
Foot	58
Pain	58
Ear	58
Pancreas	57
Aids	57
Blood	55
HIV	54

Classified term occurrences important for us

This section analyzes only queries directly from HON as they correspond better to our needs concerning patient information search. We particularly note the most frequent terms for *anatomic region, pathology, imaging modality, symptom* and *treatment*, as these are axes to model search tasks along.

Table 3 - Frequent terms regarding modality

Term	*Frequency*
Ultrasound	47
Ecg/ekg	34/32
MRI	33
X-ray	21
Endoscopy	18

Table 3 shows modalities searched for. Interestingly, a commonly used modality (CT) is not mentioned often, whereas ECG, often discarded in medical image databases, is frequently used as it corresponds to the information needs.

Table 4 - Frequent terms regarding symptoms

Term	*Frequency*
Bacteria	64
Pain	58
Burns	42
Stress	37
Blood pressure	30

Table 4 shows symptoms searched for, where symptom is taken in a broad sense. Bacteria is not a symptom but might be interpreted from patients with flu-like symptoms looking for more information on a particular situation.

Table 5 - Frequent terms regarding treatments

Term	*Frequency*
Injection	67
Surgery	46
Stethoscope	36
Anesthesia	24
Vaccination	22

Table 5 lists terms concerning treatments, taken in a wide sense, as stethoscope is not a treatment.

Table 6 - Frequent terms regarding anatomic region

Term	*Frequency*
Heart	381
Brain	211
Liver	101
Kidney	77
Lung	69

In Table 6, frequent anatomic regions are listed that correspond well to the most frequent causes of death [14]. Also the search terms regarding pathology correspond well to diseases mentioned in [14]. Only Marfan is surprisingly frequent.

The 500 most frequent terms were analyzed accounting for almost half the search terms in total. Besides the identified five axes, some other terms are frequently queried, which are hard to classify: Human body (41), smoking (38), CPR (computerized patient record, 33), cardiology (26). It is hard to know what images or videos the users were searching for.

Table 7 - Frequent terms regarding pathology

Term	*Frequency*
Asthma	242
Diabetes	160
Cancer	98
Marfan	93
Aids/HIV	57/54

Constraints to define search tasks based on the results

From the most frequent concepts and the average number of query terms it becomes clear that users express fuzzy information needs and describe them with few terms. As the information in the HON queries corresponded better to our goal, we only used these. It is clear that information needs are often broad and it seems to aim at general illustrations (CPR, human body, AIDS …) than towards precise images of a particular modality and anatomic region. Illustrations also need to be taken into account as

frequent query words such as doctor, nurse, injection or bacteria show.

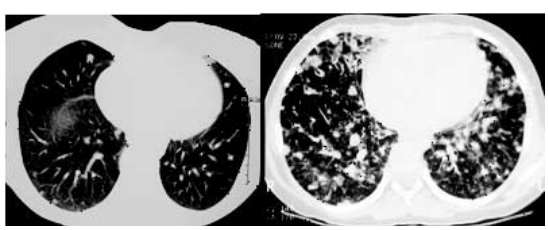
Table 8 - A collection of longer queries

Term	*Frequency*
Autonomic nervous system	16
Heart conduction system	10
Artrite reumatoide juvenil	9
Lupus vasculitis central nervous system	8
Fetal alcohol syndrome	7
Sickle cell anemia	7
Epilepsy frontal lobe	6
Respiratory distress syndrome adult	6
Spinal cord compression	6
Shoulder impingement syndrome	6

Other queries contained expected concepts but not as detailed as desired. If looking for images of the heart, all modalities, views and pathologies combined produce an extremely large number of images to be found. Such tasks are not suited to find out more about the quality of a retrieval system. For this reason, we evaluated the most frequent queries with at least three words. Table 8 lists these frequent search terms. The table shows that several terms still contain a single concept (autonomic nervous system). Most queries contain two distinct concepts, either pathology and anatomic information (epilepsy frontal lobe) or a disease and a patient group (respiratory distress syndrome adult). Still, few of these queries can be taken as query tasks for a benchmark directly.

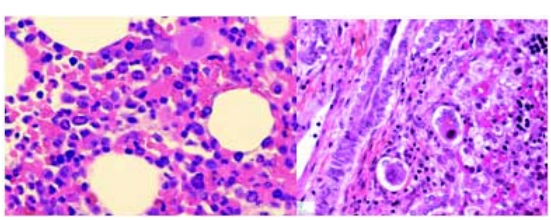
Example Query tasks of ImageCLEFmed

Finally, it was decided to use 30 real but rarer queries of the log files that cover at least two of the axes modality, anatomic region, and pathology. Example topics with example query images can be seen in Figures 1 and 2.



Show me chest CT images with nodules.
Zeige mir CT Bilder der Lunge mit Knötchen.
Montre-moi des CTs du thorax avec nodules.

Figure 1 - A visual query of ImageCLEFmed 2006



Show me microscopic images showing parvovirus infection.
Zeige mir Mikroskopien mit einer Parvovirusinfektion.
Montre-moi des images microscopiques qui montrent une infection parvovirale

Figure 2 - A semantic query for ImageCLEFmed 2006

The 30 query topics generated in this way were sent to all 40 participating research groups together with an image database. After retrieval experiments by participating groups and pooling of results, a group of physicians performed relevance judgments to compare the retrieval results of the participating retrieval systems. More about the results can be read in [15].

Discussion and conclusions

The normalization of query terms that we applied is not completely sufficient for a system that is used in several languages. A translation of the terms towards a single language or terminology would be best but with most queries being single words, this is difficult. At least 10 languages were identified. Spelling errors and abbreviations were other problems. Part of this was corrected with manual analysis but a large number of queries for the same terms could not be combined.

It can be seen that many queries for visual medical content are being performed with HONmedia search. About 52'000 queries in ten months is a large number for a small specialized search engine. Some queries are not for medical content but erotic, which is a phenomenon known by all search engines, particularly searches for images. Many queries are for illustrations of broad concepts, where the users seem to be willing to browse through a large number of varying results without a clear idea in mind and rather to illustrate an article or a presentation. Most queries are for a particular anatomic region or a certain disease. Users of the search engine do not seem to be used to supplying precise information needs concerning images. They follow the behavior of textual Internet search using broad concepts. Most image databases on the web are not well annotated and much of the information is incomplete resulting possibly in poor results.

Compared to text analysis and retrieval, medical visual information retrieval is still in its infancy. Currently, large data sets are being created and made available. Still, the applied search methods are mostly based on text, only. Techniques for visual retrieval do exist [9] and if we want to apply them in real clinical settings we need to build prototypes and make users familiar with the techniques, the possibilities and the limitations. In this sense, ImageCLEFmed is an important initiative for bringing image retrieval systems closer to routine use, through evaluating their quality. To do so, the common image databases need to be shared and realistic visual information needs have to be defined. For this, resources such as the HONmedia log files are important for us as only few medical visual search engines exist in routine use. It is also important to educate users to define their information needs more precisely using text as well as visual means and also relevance feedback.

An interesting future research topic is the analysis of query terms over short time frames. How does this behavior change with respect to events in the world (such as the bird flu)? Could the beginning of a flu outbreak be detected through keyword changes for related terms? Medical

image search on the Internet and in institutional databases has a high potential but more research is needed and particularly prototypes that can be made available to the users for testing to find out more about concrete information needs.

Acknowledgements

This work was partially supported by the Swiss National Science foundation with grant 205321-109304 and by Grant 0325160 of the US National Science Foundation..

References

[1] Rice RE, Influences, usage, and outcome of Internet health information searching: Multivariate results from the Pew surveys, International Journal of Medical Informatics, 75:8-28, 2006.

[2] Gaudinat A, Boyer C, WRAPIN (Worldwide online Reliable Advice to Patients and Individuals). In MEDNET 2003, The 8th Annual World Congress on the Internet and Medicine, Geneva, Switzerland.

[3] Hersh WR, Hickam DH, How well do physicians use electronic information retrieval systems? A framework for investigation and systematic review, Journal of the American Medical Association, 1998, 280: 1347-1352

[4] Gould P, The rise and rise of medical imaging, physicsweb, 16(**8**), August 2003.

[5] Smeulders AWM, Worring M, Santini S, Gupta A, Jain R, Content-Based Image Retrieval at the End of the Early Years, IEEE Transactions on Pattern Analysis and Machine Intelligence 22(**12**) pp 1349-1380, 2000.

[6] Markkula M, Sorminen E, Searching for photos - Journalists' practices in pictorial IR. In JP Eakins, DJ Harper, J Jose (Eds.), The Challenge of Image Retrieval Newcastle upon Tyne, UK, 1998.

[7] Choi Y, Rasmussen EM, Users' relevance criteria in image retrieval in American history, Information Processing and Management 38: 695-726, 2002.

[8] Tagare HD, Jaffe C, Duncan J, Medical Image Databases: A Content-Based Retrieval Approach, Journal of the American Medical Informatics Association (JAMIA), 4(**3**):184-198, 1997.

[9] Müller H, Michoux N, Bandon D, Geissbuhler A, A review of content-based image retrieval systems in medicine – clinical benefits and future directions, International Journal of Medical Informatics, 73, pp 1-23, 2004.

[10] Clough P, Müller H, Deselaers T, Grubinger M, Lehmann T, Jensen J, Hersh W, The CLEF 2005 Cross-Language Image Retrieval Track, Springer Lecture Notes in Computer Science LNCS **4022**, pp 535-557, 2005.

[11] Hersh W, Müller H, Gorman P, Jensen J, Task Analysis for Evaluating Image Retrieval Systems in the ImageCLEF Biomedical Image Retrieval Task, Slice of Life, multimedia in medical education, Portland, OR, USA, 2005.

[12] Müller H, Despond-Gros C, Hersh W, Jensen J, Lovis C, Geissbuhler A, Medical professionals' image search and use behavior, Medical Informatics Europe, 2006.

[13] Paliouras G, Papatheodorou C, Karkaletsis V, Spyropoulos CD, Tzitziras P, "From Web Usage Statistics to Web Usage Analysis," IEEE Conference on Systems Man and Cybernetics, 1999.

[14] Anderson RN, Smith BL, Deaths: Leading causes for 2001, National Vital Statistics Report 52(**9**):1-86, 2003.

[15] Müller H, Deselaers T, Lehmann T, Clough P, Kim E, Hersh W, Overview of the ImageCLEFmed 2006 Medical Retrieval and Medical Annotation Tasks, Springer Lecture Notes in Computer Science (LNCS), 2007 – to appear.

Address for correspondence

Henning Müller (PhD)
University & Hospitals of Geneva, Medical Informatics Service
24, rue Micheli-du-Crest, 1211 Geneva 14, Switzerland
Tel +41 22 372-6175, Fax +41 22 372-8680
henning.mueller@sim.hcuge.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Improving Computer Aided Disease Detection Using Knowledge of Disease Appearance

Tatjana Zrimec[a,b] James S. Wong[b,]

[a] *Centre for Heath Informatics, University of New South Wales, Australia*
[b] *School of Computer Science & Engineering, University of New South Wales, Australia*

Abstract

The accurate quantification of disease patterns in medical images allows radiologists to track the progress of a disease. Various computer vision techniques are able to automatically detect different patterns that appear on images. However, classical pattern detection approaches do not perform satisfactorily on medical images. The problem is that texture descriptors, alone, do not capture information that is pertinent to medical images, i.e. the disease appearance and distribution. We present a method that uses knowledge of anatomy and specialised knowledge about disease appearance to improve computer-aided detection. The system has been tested on detecting honeycombing - a diffuse lung disease pattern in HRCT images of the lung. The results show that the proposed knowledge guided approach improves the accuracy of honeycombing detection. A paired t-test, shows the improvement in accuracy to be statistically significant ($p < 0.0001$).

Keywords:

lung HRCT, lung diagnosis, computer aided diagnosis, medical imaging, honeycombing detection

Introduction

Medical imaging systems are constantly improving in image quality because of increased image resolution. This results in a growing number of images that have to be inspected for diagnosis. For example, high resolution CT (HRCT) imaging protocols of the lungs can generate from 40 up to 600 images per study. These high-resolution axial images provide anatomic detail similar to that available from gross pathology specimens of lung slices [1]. Now radiologists can clearly see the alterations in lung anatomy caused by a disease process. Unfortunately, image analysis is still performed manually, which is often a difficult and time-consuming task. Consequently, there is an increasing need for computerised image analysis to facilitate image-based diagnosis.

We are developing a system for computer-aided detection of diffuse lung diseases, a large group of disorders that primarily affects the lung parenchyma. They are characterised by specific abnormal findings, mostly texture-like in appearance. Consequently, most of the automated detection algorithms, being developed to analyse CT scans are texture based. The classical approach is to use a set of image features to describe the image content and to use some classification scheme to distinguish between different patterns. For example, Uppaluri et al. [2] used twenty-two independent texture features to characterise a tissue pattern in the overlapping square regions of the lung. A Bayesian classifier was trained to discriminate between six different patterns. Uchiyama et al. [3] proposed a similar texture based technique. They trained an Artificial Neural Network with twelve features, calculated on regions with different sizes, to classify new regions. The system was trained to distinguishing between seven different patterns, which included normals and six patterns associated with diffuse lung diseases.

Our system, developed to automatically detect lung disease patterns, adopts a similar approach. However, we use a much bigger set of image attributes to describe the content of the image. We experimented with different attributes subsets and different learning schemes to improve the system's performance. The results reveal that classical pattern detection approaches do not perform satisfactorily on medical images. The problem is that texture descriptors, alone, do not capture information that is pertinent to medical images, i.e. the disease appearance and distribution. Therefore we incorporated domain knowledge of lung anatomy and lung structure to help and improve image analysis.

In this paper we focus on detecting honeycombing, an important diffuse lung disease pattern, in HRCT images of the lung. As the goal of the system is to provide radiologists with a second opinion on a lung diagnosis, it is important to achieve high accuracy. In this paper we present a new method developed to improve computer-aided detection. The method uses specialised knowledge of disease appearance in axial images. It also uses information about lung regions that are used in radiology reporting [9]. To determine if using knowledge can significantly improve the system's performance, we incorporated the knowledge-guided approach into two classification methods, one based on decision tree learning and the other using Naïve Bayes.

In the remainder of the paper, we present a computer-aided detection system using a classical pattern detection

approach. We then present the improved system and compare their performance.

Materials and methods

Lung diseases – honeycombing pattern

Honeycombing is one of the main indicators of diffuse lung diseases. It can be seen in many diseases leading to end-stage pulmonary fibrosis. Honeycombing is characterised by small, uniform (2-10mm) cystic air spaces with well-defined thick walls, (See Figure 1 left). Honeycomb cysts usually form clusters that have the characteristic appearance of "honeycombing" on HRCT images. The visual appearance of honeycombing in cross-section scans is a combination of dark and light patches. It is one of the more difficult disease patterns to detect because honeycombing can often be mistaken for other normal structures in the lung, for example, bronchi and pulmonary vessels (see Figure 1 right).

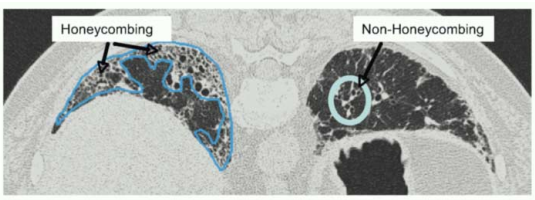


Figure 1 - Left lung- outlined region with honeycombing. Right lung –outlined example of bronchovascular structures, which has similar appearance as honeycombing.

The HRCT images

The HRCT images used for developing the computer aided disease detection system were obtained from a radiology practice. They were acquired using a SIEMENS scanner and a standard imaging protocol with 750ms exposure time. The HRCT generates volume data with spatial resolution 512x512 and 1.0mm slice thickness. For diffuse lung disease detection, radiologists usually use data with 10mm or 15mm slice gap, The data are stored in DICOM format as 16-bit greyscale images.

Describing the disease

There are specialised texts that describe how to interpret HRCT images, accompanied by illustrative examples [1]. Although this highly specialised knowledge is very useful for humans, computers cannot make direct use of it.

For computers to recognise a disease in medical images, the content of the image is represented by a set of image attributes, calculated for regions of interest (ROI) in an image. The values of these attributes depend on the characteristics of the regions. We can use these characteristic values to distinguish between normal and pathological regions as well as between different pathologies.

The method

The easiest way for a radiologist to communicate expert knowledge about how diseases appear in HRCT images is to provide examples. Using a specially developed image mark up system [4], we collected a set of images with marked and labelled regions of honeycombing and other lung diseases patterns. Figure 2 (top) shows example of an HRCT image with marked regions. Having examples of image regions with and without a disease, we were able to use supervised machine learning to generate rules for recognising different patterns in HRCT images.

The method consists of there main steps:

1. Data preparation:
   a) Image pre-processing and segmentation
   b) Feature extraction - calculating attributes for regions of interest
2. Knowledge generation - training:
   a) Feature selection – finding informative attributes for a particular disease pattern
   b) Generating rules via machine learning
3. Knowledge verification -testing the quality of the learned rules:
   a) On part of the training data
   b) On new data

Data pre-processing and segmentation

As we are interested in detecting patterns in the lungs, we first pre-process the images and segment the lungs. We have developed a lung segmentation technique based on adaptive thresholding, morphological operators and active contour snakes. Adaptive thresholding is applied to segment the darker regions in the image that represent the air-filled lung. Morphological operators are then used to include structures within the lung that have a high attenuation. Active contour snakes [5] are used to generate the lung contours (see Figure 2 bottom row).

Feature extraction

Having segmented the lung, we proceed to extract features from the image that best represent the underlying texture. A set of attributes was calculated for each pixel and it's surrounding area. We used a ROI with two window sizes, 7x7 and 15x15 pixels, to capture the characteristics of small and larger honeycombing cysts.

We calculate first and second order texture attributes and grey-level difference for each ROI [6]. The first order texture attributes measure the grey-level distribution within the ROI. Those attributes include: the mean HU[1], variance, skewness, kurtosis, energy and entropy. The second order features describe the spatial distribution of the grey-levels within these ROIs. To do this, a co-occurrence matrix is calculated that specifies the frequency of a particular grey-level occurring near another grey-level. Each pixel, with it's surrounding area, is represented by 63 attributes per window, resulting in a feature vector with 126 attributes, (63 for ROI_{7x7} and 63 for ROI_{15X15}).

1 Hounsfield unit (HU) a unit used in medical imaging (CT or MRI scanning) to describe the amount of x-ray attenuation of each "voxel" (volume element) in the three-dimensional image.

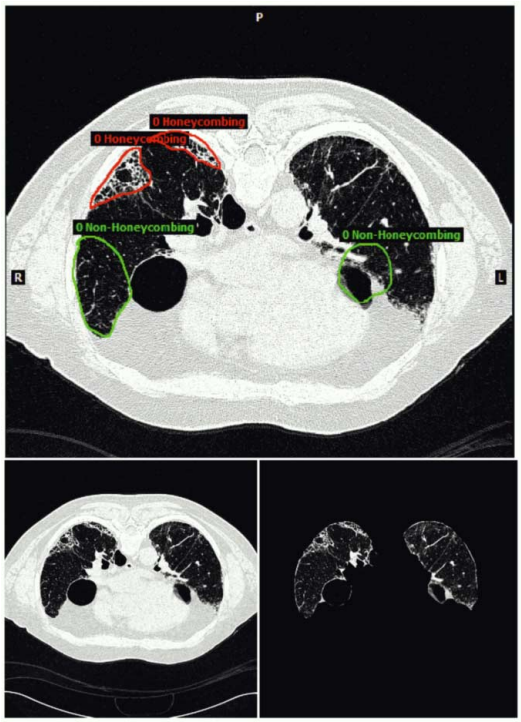


Figure 2 - Top: an image with marked and labelled regions by a radiologist showing disease patterns; bottom left: original HRCT image, bottom right: segmented lungs.

Knowledge generation process - feature selection

With a feature vector of 126 dimensions, the classifier generated would be computationally intractable. We reduce the dimensionality by selecting a subset of features that best discriminates honeycombed and non-honeycombed regions. In this study, we use Correlation-based Feature Selection (CFS) [7]. CFS selects subsets of attributes that are highly correlated with the class and that have low intercorrelation.

Generating rules via machine learning

Even with the reduced set of attributes, it is difficult to manually determine the attribute values that are characteristic for a particular pattern. We used supervised machine learning to automatically generate rules for discriminating between different patterns. In supervised learning, a set of pre-classified training examples is used to generate classification rules. In our case, the training examples consist of a set of attribute values representing a region with or without honeycombing pattern. The training set was prepared from the images with labelled regions provided by radiologists (see Figure 2 top).

We experimented with two machine learning algorithms to build the classifier: Naive Bayes and the decision tree learner, J48, both implemented in Weka data mining environment [8]. The Naive Bayes algorithm is based on a probability model. The probability of a class given a feature vector is determined using Bayes' rule:

$$P(c \mid F) = \frac{P(F \mid c)P(c)}{P(F)} \quad (1)$$

where c is the class and F is the vector of features. The class with the highest probability is assigned to the ROI. Although the Naive Bayes approach is optimal when the features are independent, in reality it still works well without this assumption. The decision tree learner generates a tree in which nodes represent tests on attributes, branches represent attribute values and leaf nodes represent classes, i.e. decisions. More informative attributes appear higher in the tree.

To train the system to detect honeycombing pattern, we used 42 images from 8 different patients that showed patterns representative of honeycombing and non-honeycombing tissue. After prepossessing and segmentation, feature extraction and selection were performed as described previously. A set of 18,407 labelled region of interest were used for training the machine learning algorithms from which 9,467 ROIs contained honeycombing and 8,940 ROIs did not. Two classifiers were built, one using Naïve Bayes and the second using decision tree induction.

Knowledge verification

Ten-fold cross validation was used to estimate the accuracy of the learned classifiers. In this validation scheme, 90% of the training data are randomly sampled for learning and 10% for testing. This is repeated 10 times and the results averaged.

The performance of the classification during training and testing was evaluated by calculating *accuracy*, *sensitivity, and specificity*. In our case *accuracy* measures the proportion of the lung that is classified correctly. *Sensitivity* determines the proportion of actual honeycombing that has been detected as honeycombing. *Specificity* measures the amount of non-honeycombing that has been classified as non-honeycombing.

Although the *accuracy*, *sensitivity* and *specificity* were comparable with the results published in the literature, we were not satisfied with the system's performance. It produced some spurious honeycombing classifications in regions where honeycombing cannot appear.

Improvements based on domain knowledge

Instead of developing post-processing methods for handling misclassifications, we decided to make use of domain knowledge about the lung structure as well as expert knowledge about the appearance of diseases. For example, Web [1] pp 91, states that "Honeycombing results in cysts …which have a peripheral predominance". This simple statement is not simple to implement. We first had to develop a model of the human lung. Next we had to develop algorithms that use anatomical knowledge to automatically generate lung regions, such as, peripheral, central, apical and basal, which are frequently used in disease reporting [9], [10]. These enabled us to determine the

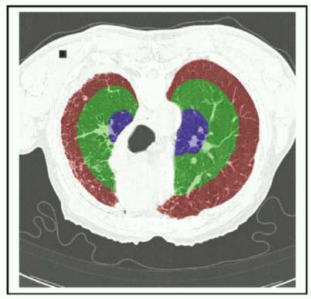


Figure 3 - An HRCT image with lung regions: blue-central, red - peripheral.

lung periphery on each axial scan, which helped in desease classification.

Knowledge guided classification

In many systems (e.g. [2, 3]), all regions within the lung are classified starting from the top of the image. However, for diseases that show honeycombing, the pattern spreads from the periphery of the lung. We developed a knowledge-guided strategy for classification. This strategy uses seeded region-growing [11] and works as follows:

- The algorithm initially only classifies peripheral regions. Peripheral regions are determined using the lung regions masks. ROIs in the periphery of the lung with honeycombing are set as the "seed points" for the algorithm.
- The algorithm only classifies a region of interest if it is near other ROIs already classified as honeycombing.
- The algorithm will stop when there are no more ROIs to consider.

In summary, the knowledge guides the system to classify all ROIs that are either in the periphery of the lung or in close proximity to other ROIs classified as honeycombing.

Results

In order to test the clinical viability of the system, we evaluated the performance of our system on part of the training data and on new, previously unseen data.

Testing on the training data **-** We used ten-fold cross validation. The number of ROIs used for testing varies for each fold, as the size of the lung in each slice affects the number of ROIs that we extract. On average 9,337 ROIs were used for testing (876 ROIs for honeycombing and 8,461 ROIs for non-honeycombing). The number of ROIs containing non-honeycombing was significantly larger as most of the lung region does not show honeycombing. The average of the results are shown in Table 1.

From the results presented in Table 1, it can be seen that the knowledge-guided approach improved the accuracy of the honeycombing detection. The improvement in accuracy is attributed to a decrease in false positive classifications (sometimes by over 2%). The increase in accuracy shows that the technique is well suited for honeycombing detection. A paired t-test, shows the improvement in accuracy to be statistically significant ($p < 0.0001$).

Table 1 – Results of the two classifies: Decision Tree Induction J48 - (DTI-J48) and Naïve Bayes used with classical and knowledge-guided approach.

Classifier	DTI-J48	DTI-J48	Naïve Bayes	Naïve Bayes
	Classical	Knowledge-guided	Classical	Knowledge-guided
Accuracy	88.20%	89.70%	85.50%	87.20%
Sensitivity	96.70%	96.60%	97.50%	97.40%
Specificity	86.80%	88.60%	83.50%	85.50%

Testing on new, previously unseen data **-** Images from 8 patients, 4 patients with homecoming present and 4 patients with different disease patterns were used to test the performance of the detection system. The evaluation was performed on six images with honeycombing present and six images with out honeycombing. In total, there were 3150 regions with honeycombing and 60318 regions without honeycombing. Sensitivity of the algorithm dropped to 85%.

Conclusion

The accurate quantification of disease patterns in medical images allows radiologists to track the progress of a disease. We have developed a system that uses machine learning to automatically detect honeycombing patterns in HRCT images of the lungs. Applying a classical, texture-based, approach resulted in over detection of honeycombing. It was not possible with simple post-processing to remove the false positive regions. We improved the perfor-

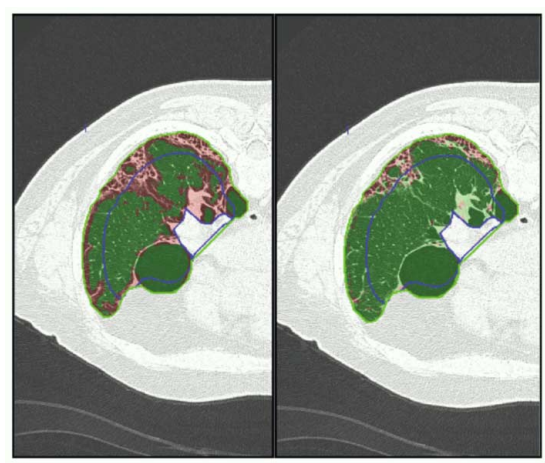


Figure 4 - Left: classification without using knowledge; right: knowledge-guided classification; red regions are classified as honeycombing; the blue line is the mask defining the periphery.

mance of the system by using knowledge-guided classification.

We tested the system with part of the training data and with new, previously unseen data. The results showed a high degree of accuracy (89.7%) and sensitivity (96.6%) on the training data. The accuracy sensitivity, however drop from 97 to 71% when testing on new data.

Experiments in building classifiers with different machine learning algorithms, Naïve Bayes and J48 decision tree learner, showed that the knowledge-guided classification performs better in both cases. The results showed that using knowledge-guided classification using the texture based Naïve Bayes classification lead to significant improvement according to pared t-test.

The domain knowledge not only improved the results of the classification, but it also improved the representation of the results. For example, a computer aided system without knowledge of lung structure reports that 5% of the fifth image contains honeycombing and that 20 % of the images 9, 10 and 11, also contain homecoming. Our system reports that 5% of honeycombing was detected in the apical area and 20% in the basal area, predominantly in the lower lobe of the left lung. Being able to use knowledge of lung anatomy in image analyses will significantly improve the detection and quantification of other lung diseases.

Acknowledgments

We thank Dr. Peter Wilson for providing medical knowledge and guidance and Medical Imaging Australia for providing data. We thank Claude Sammut for his comments. This research was supported by the Australian Research Council.

References

[1] Webb WR, Müller NL, Naidich DP. High-Resolution CT of the Lung. 3rd ed. Lippincott Williams & Wilkins, Philadelphia, 2001.

[2] Uppaluri R, Hoffman EA, Sonka M, Hartley PG. Computer Recognition of Regional Lung Disease Patterns. American Journal of Respiratory and Critical Care Medicine 1999:160(2): 648–654.

[3] Uchiyama Y, Katsuragawa S, Abe H, Shiraishi J, Li F, Li Q, Zhang CT, Suzuki K, Doi K. Quantitative computerized analysis of diffuse lung disease in high-resolution computed tomography. Medical Physics 2003:30(9): 2440–2454.

[4] Rudrapatna M, Sowmya A, Zrimec T, Wilson P, Kossoff G, Wong J, Busayarat S, Misra A, Lucas P, LMIK – Learning Medical Image Knowledge: An Internet-based medical image knowledge acquisition framework, Electronic Imaging Science and Technology - IS&T/SPIE's 16th Annual Symposium, San Jose, CA. 2004.

[5] Papasoulis J, LMIK - Anatomy and Lung Measurements using Active Contour Snakes, Undergraduate thesis, supervisor: T. Zrimec, Computer Science and Engineering, University of New South Wales, Sydney, Australia, 2003.

[6] Haralick RM, "Statistical and Structural Approaches to Texture," Proceedings of the IEEE 67, pp. 786– 804, 1979.

[7] Hall, MA. Correlation-based feature selection for discrete and numeric class machine learning. In: ICML '00: Proceedings of the Seventeenth International Conference on Machine Learning, San Francisco, CA, USA, Morgan Kaufmann Publishers Inc. 2000; pp. 359–366

[8] Witten, IH, Frank E. Data Mining: Practical machine learning tools and techniques. 2nd edition, Morgan Kaufmann, San Francisco, 2005.

[9] Zrimec T, Busayarat S, Wilson P. A 3D model of the human lung with lung regions characterization. ICIP 2004 Proc. IEEE Int. Conf. on Image Processing, 2004, pp. 1149–1152.

[10] Zrimec T, Busayarat S, 3D modelling and visualisation of the human lung, Proc. 2nd Int. Symp. on 3D Data Processing, Visualization and Transmission, Thessaloniki, Greece, September 2004.

[11] Adams R, Bischof L. Seeded region growing. Pattern Analysis and Machine Intelligence, IEEE Transactions on 1994: 16(6): 641–647.

Address for correspondence

Tatjana Zrimec
Centre for Health Informatics
University of New South Wales
Sydney, 2052
tatjana@cse.unsw.edu.au
Tel: +61 2 9385 9034
Fax: +61 2 9385 9006

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

MR Atlas for Articular Cartilage Morphology: Potential to Detect Shape Differences

Hussain Z. Tameem[a,b], Usha S. Sinha[b]

[a] *Department of Biomedical Engineering, University of California, Los Angeles, United States*
[b] *Medical Imaging Infomatics Group, University of California, Los Angeles, United States*

Abstract

An atlas of the cartilage was created using free form transformation of MR images of the cartilage from 20 subjects. The deformation required to move each voxel to its corresponding location in the atlas is used to determine the differences in shape between cartilages of subjects in a population. Based on these active shape models, it is possible to localize regions of high morphological variance in population cohorts. The atlas, reported here, is based on 20 male subjects; ten symptomatic of arthritis and ten asymptomatic. The active shape models based on this atlas show regions of high morphological variance corresponding to cartilage thinning in the arthritic group. This method has the potential to differentiate between normal and arthritic population groups by detecting subtle morphological changes in articular cartilage.

Keywords:

active shape models, articular cartilage, image registration, magnetic resonance imaging, morphometry

Introduction

Osteoarthritis (OA) is a complex, progressive disease of the joint characterized by degenerative and regenerative morphological and structural changes in the articular cartilage and subchondral bone [1]. OA is a slowly progressing disease characterized clinically by pain, enlargement and deformity of the joints, and limitation of the motion. OA is the most prevalent form of arthritis and leading cause of disability and work limitation among adults resulting in enormous cost to society [2, 3]. Approximately 21 million American adults have physician diagnosed OA, [4] a diagnosis usually based on the combination of joint symptoms and radiographic changes. The prevalence of OA in a population is difficult to determine because: 1) the degree of radiological changes in symptomatic individuals varies greatly and 2) many individuals with radiographic evidence of OA have no symptoms. By age 60 nearly half of the population has radiographic evidence of OA in one or more joints, and by age 80 these findings are universal [5, 6].

Among the various sites being affected in OA, knee is the major source of reported disability and loss of function. About 40% of the adult population age 55 and older has frequent knee pain or definite x-ray evidence of knee OA [7-9]. Advanced OA accounts for majority of knee replacements surgeries among Medicare recipients. Well over 200,000 knee replacement procedures for OA are performed every year in the United States [2].

As of today, there are no reports of any disease modifying therapies for knee OA and all treatments are predominantly designed to relieve pain [10]. Approaches to prevent knee OA development, progression, or related disability are also very limited, in large part due to incomplete knowledge of potentially modifiable factors responsible for these outcomes.

In this paper, we report a method based on free-form transformation to generate an average shape atlas of the femoral cartilage and apply it to study shape differences in a population cohort. This method has potential applications in the detection of subtle shape differences in normal and diseased population groups.

Background

Previous research on cartilage morphological assessment includes its volume and thickness measurements and the impact of various factors on normal knee, such as sex, body weight and height, maturity and age, body mass index, leg length and foot size, knee bone size, bone mineral density, muscle mass, level of physical exercise and genetics [11,12,13,14]. But none of the research to date focuses on studying the shape changes in cartilage between normal and diseased state population. Also there are studies which indicate the use of mathematical frameworks such as principal component analysis (PCA) to describe general shape variations. Marcus [15] used PCA to study the variation in the skull measurements of rodent and bird species. The resulting principal modes were interpreted as size and gross shape components. Cootes et al. [16] applied the theory of PCA to build statistical shape models of organs based on manually chosen landmarks. This model provided the average positions of the points and the principal modes of variations were computed from the dataset. The ability of the method to locate structures in medical images was demonstrated in a set of experiments with echocardiograms, brain ventricle tracking and prostate segmentation. Le Briquer and Gee [17] applied PCA to analyze the displacement fields obtained from registering a reference image volume of the brain to a set of subjects, based on the elastic matching framework. The

analysis provided the inference of morphological variability within a population and was the basis for the construction of a statistical model for brain shape, which could be used as prior information to guide the registration process. Duchesne et al. proposed shape models for segmentation of the medial temporal brain structures [18].

Materials and methods

Image acquisition

We obtained the images from a pilot study conducted for the National Institute of Health OA initiative (OAI) version 0.A.1. MR images were acquired using a water-excitation double echo steady-state (DESS) imaging protocol with sagittal slices at 3.0T (Magnetom Trio®, Siemens). The imaging parameters for the sequence were: TR/TE: 16.3/4.7 ms, matrix: 384∞384, FOV: 140 mm, slice thickness: 0.7 mm, x/y resolution: 0.365/0.365 mm. Figure 1 shows a sagittal slice of the magnetic resonance image obtained using DESS sequence.

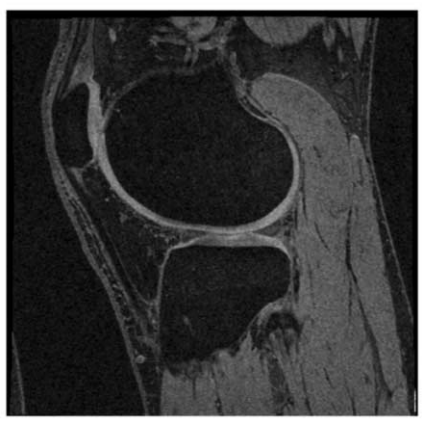


Figure 1 - MR image acquired with a DESS sequence with water-excitation at 3.0 T. The images are obtained from National Institutes of Health (NIH) OA initiative

The OAI data consists of a stratified random sample of 200 participants based on gender, sub-cohort assignment (progression and incidence) and clinic (four recruitment centers). The progression sub-cohort contains participants with symptomatic knee OA at baseline where symptoms are pain, aching or stiffness in or around the knee on most days for at least one month during the past 12 months. The incidence sub-cohort contains participants with no symptomatic knee OA at baseline, but has characteristics that place them at the risk for developing symptomatic knee OA during the study. For this study we randomly chose 20 male participants, 10 each from progression and incidence cohort group.

Atlas creation

We manually segmented the cartilage for all 20 subjects and then interpolated the raw data to a pixel resolution of 0.365x0.365x0.365 mm^3. We created the atlas in the following steps, which are illustrated in figure 2.

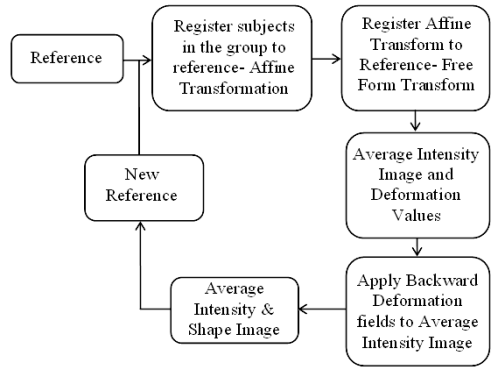


Figure 2 - Schematics of steps involved in atlas creation

Step I: We randomly chose one subject from the pool of 20 to serve as a reference to which the rest of the images were aligned utilizing the mutual information based affine transformation. Affine transform corrects for subject positioning and global size differences.

Step II: An elastic registration based on demons algorithm [19] was employed to locally map all the images in the group of subjects to the reference image using the affine transformation parameters as an initial estimation. This provides 3D deformation fields that can map the spatial locations on an individual in the group to the coordinate system of the reference. The registration algorithm computes the transformation iteratively using equation 1.

$$v_{n+1} = G_\sigma \otimes \left(v_n + G_\sigma \otimes \frac{1}{2}\left[\frac{C(T-S)\|\nabla S\|\nabla T\|}{(\|\nabla T\|^2 + \|\nabla S\|^2)(\|\nabla S\|^2 + \|\nabla T\|^2 + 2(T-S)^2)} \nabla S \right] \right) \quad (1)$$

v_{n+1} is the correction vector field at iteration n+1, G_σ is the Gaussian filter with variance σ2, ⊗ denotes the convolution, C is the scaling factor and T and S are the target and transformed images respectively. The algorithm estimates the displacement which maps a voxel at location (*x,y,z*) in *T* to the corresponding anatomical location in *S.* The algorithm is implemented hierarchically and to preserve the morphology, deformation vector fields were computed utilizing both the forward and backward transformation.

***Step III*:** A mean intensity image with the shape of the reference image is created by averaging the globally and locally transformed images of the group.

Step IV: A mean deformation field that encodes the shape variation between the reference image and average shape of the elements in the subject group is created by averaging over 3D deformation vector fields of the individual subjects of the group.

Step V: Inverse average deformation field is applied to the average intensity image to generate and average intensity and deformation image template for the group under study.

Step VI: Steps 1-5 are iterated until no significant change in the deformation field is observed relative to the previous computation. At the end of each iteration the original reference image is replaced by the average template

Figure 3 - Accuracy of registration. Left to Right: reference, test, result of affine transform and free form deformation image. Top and bottom row showing slices at different locations. Outline from reference image shown superposed on test & aligned images.

constructed at *Step V* generating both average shape (morphometric) and intensity atlases that represent the centroid of the population data set.

Active shape models

Active shape models are used to represent the variance in cartilage shape within a given population. Active shape models based on principal component analysis of the deformation fields were created using the data from the last iteration of the atlas creation procedure [17]. With reference to the atlas creation, each iterative process results in an average atlas and deformation field d_{ir}. The deformation field, $d_{ir,}$ is the amount required to move the voxel from its original position (after global affine transformation) to the corresponding location in the atlas. Here *r* refers to the voxel, *i* represent the subjects (1-20 used to create the atlas) and *d* represents the deformation vector. The following analysis based on principal component analysis for data reduction was performed on all n voxel of the cartilage and consists of the following steps.

Step I: Calculation of the mean deformation for N subjects at each voxel as shown in equation 2 where dmean is the mean deformation at any voxel over all subjects.

$$d_{mean} = (1/N)\,(\Sigma\, d_i) \tag{2}$$

Step II: Computation of the deviation from the mean value as shown in equation 3

$$\Delta d_i = d_i - d_{mean} \tag{3}$$

Step III: Calculation of n∞n covariance matrix, C, to find the basis for the space as shown in equation 4

$$C = (1/N)\,(\Sigma \Delta d_i d_i^T) \tag{4}$$

Step IV: Diagonalization of the covariance matrix to obtain the eigenvectors, ν_k and the eigenvalues, λ_k.

Step V: Construction of the linear model as shown in equation 5, where $\nu = (\nu_1, \nu_2, \ldots, \nu_k)$ is the matrix of the first eigenvectors, and W_s is a vector of weights, also called the shape coefficient. This results in a shape model.

$$d = d_{mean} + \nu W_s \tag{5}$$

The shape variations by ±2 SD from the mean shape along the first two principal modes were generated using eigenvalues and eigenvectors derived from this analysis. These images were synthesized by setting the weights of the first or the second modes at $\pm\sqrt{\lambda_i}$ and $\pm 2\sqrt{\lambda_i}$ (where i = 1, 2 for the first and second eigen mode respectively) and all other weights to zero. The synthesized images provide a visual representation of the possible variance in shape of the cartilage based on 20 image sets to create the atlas.

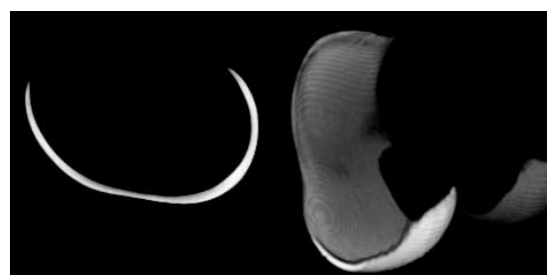


Figure 4 - Left: Shows the sharp edges on the atlas. Right: 3D volume reconstruction of the cartilage from the atlas

Results

The alignment of test image to the reference image using affine and freeform transformation is shown in figure 3. A region of interest is selected on the reference image and overlaid on the test, affine and the image obtained after free form transformation to show the accuracy of registration. It can be clearly seen that the affine transform corrects for positional and global scale changes and the local deformation corrects for physiological changes. The accuracy of the alignment can be visually evaluated by the sharp edges of the cartilage in the 2D images of the atlas as

Figure 5 - 3D active shape models. Average shape shown in the middle top row with standard deviation (SD) variations along the first and second mode shown in rows 2 and 3: Left to right - 2SD, -1 SD, +1 SD and +2 SD. Arrows indicate regions with changes from average bone shape

shown in figure 4. At the end of each iteration the reference image is replaced by the average image which moves closer to the centroid of the images in the group. There is no significant difference in the atlas after the third iteration confirming its convergence. Figure 4 shows 3 dimensional rendering of the cartilage generated from the atlas volume.

Figure 5 shows the active shape models and the variations in shape for ±2SD along the two leading eigenmodes. The outline of the cartilage is overlaid on the variations along the first and second modes that show the variations seen in the current set of 20 subjects. The first mode shows the larger changes since it captures the largest variation in the data as compared to the second mode.

Discussion

We successfully developed an atlas for the articular cartilage derived from the MR images at 3T. This is the first report of creating a cartilage atlas from images acquired at high resolution and isotropic resolution (0.365x0.365x0.365). The results show very accurate alignment which could be used for clinical purposes. We are currently working on creating a 3D Active contour without edges segmentation algorithm proposed by Chang T and Vesse L to extract the cartilage from the MR images.

We hypothesize that it is possible to automatically determine the cartilage location by analyzing its overall shape variation. If this hypothesis holds, then we can reverse the process and use the information obtained from the shape analysis to automatically segment cartilage from rest of the structures. In future we intend to use the active shape models and show that unsupervised learning can be used to explore the anatomy and facilitate segmentation. This methodology could potentially be used to classify different population groups. Structural shape characterization using PCA has been used to study gender and disease-related morphological differences in the corpus callosum, putamen, ventricles and hippocampus [20, 21]. It should be noted that within the scope of this paper we demonstrate the feasibility of generating the active shape models and that the application for classification will require far more image volumes to be included in the training set from the population cohorts of subjects under investigation.

Conclusion

We have developed an atlas for the cartilage and active shape models which when combined can be used to detect the subtle shape changes in cartilages. We see significant changes using this technique within the group of 20 subjects we selected. These models have a potential to be used in the future to discriminate normal and diseased states with larger databases.

Acknowledgments

We would like to thank The OAI study sponsored by NIH for providing the data used in this study

References

[1] Aubray JH, Jr, Sokoloff L 1989 Pathology of osteoarthritis. In: McCarty DJ (ed.) Arthritis and Allied Conditions, 11th Ed. Lea & Febiger, Philadelphia, PA, USA., pp. 1571-1580.

[2] Dunlop DD, Manheim LM, Yelin EH, Song J, Chang RW. The cost of arthritis. Arthritis Rheum 2003; 49(1):101-13.

[3] Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, Kelly-Hayes M, Wolf PA, Kreger BE, Kannel WB. The effects of specific medical conditions on the functional limitations of elders in the Framingham study. Am J Public Health 1994;84:351-8.

[4] Lawrence RC, Helmick CJ, Arnett FC, Deyo RA, Felson DT, Giannini EH. Estimates of the prevalence of arthritis

and selected musculoskeletal disorders in the United States. Arthritis Rheum 1998; 41:778-99.

[5] Bagge E, Bjelle E, Valkenburg HA, Svanborg A. Prevalence of radiographic osteoarthritis in two elderly european populations. Rheumatol Int. 1992;12:33-8.

[6] Van Sasse JLCM, Van Romunde LKJ, Cats A, Vandenbroucke JP, Valkenburg HA. Epidimealogy of osteartritis: Zoetermeer survey. Comparison of radiological osteoarthritis in a Dutch population with that in 10 other population. Ann Rheum Dis 1989;48:271-80.

[7] Urwin M, Symmons D, Allison T, Brammah T, Busby H, Roxby M, Simmons A, Williams G. Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation. Annals of Rheumatic Diseases 1998; 57:649-55.

[8] Peat G, McCarney R, Croft P. Knee pain and Osteoarthritis in older adults: a review of community burden and current use of primary health care. Annals of Rheumatic Diseases 2001; 60:91-7.

[9] Felson DT, Zhang Y. An update on Epidemiology of knee and hip osteoarthritis with a view to prevention. Arthritis Rheum 1998;41(8):1343-55

[10] Felson DT, Lawrence RC, Hochberg MC, McAlindon T, Dieppe PA, Minor MA, Blair SN, Berman BM, Fries JF, Weinberger M, Lorig KR, Jacobs JJ, Goldberg V. Osteoarthritis: New Insights. Part 2: Treatment Approaches. Ann Internal Med 2000; 133(9): 726-37

[11] Cicuttini F, Forbes A, Morris K, Darling S, Bailey M, Stcukey S. Gender differences in knee joint cartilage volume as measured by magnetic resonance imaging. Osteoarthritis Cartilage 1999; 7:265-71.

[12] Ding C, Cicuttini F Scott F, Glisson M, Jones G. Sex differences in knee cartilage volume in adults: role of body and bone size, age and physical activity. Rheumatology (Oxford) 2003.

[13] Cicuttini FM, Teichtahl AJ, Wluka AE, Davis S, Strauss BJ, Ebeling PR. The relationship between body composition and knee cartilage volume in middle-aged subjects. Arthritis Rheum 2005;52:461-7.

[14] Jones G, Ding C, Scot F, Cicuttini F. Genetic Mechanism of knee osteoarthritis: a population based case-control study. Ann Rheum Dis 2004; 63:1255-9.

[15] Marcus L. Traditional Morphometrics In: Rohlf FJ, Bookstein FL, editors. Proceedings of the Michigan Morphometrics Workshop, The University of Michigan Museum of Zoology, Ann Arbor, 1990. p. 77-122.

[16] Cootes T, Hill A, Taylor CJ, Haslam J. Use of active shape models for locating structures in medical images. Image Vision Comput 1994; 12548:126-32.

[17] Le Briquer L, Gee JC. Design of a statistical model of brain shape. Lect Notes Comput Sci 1997; 1230:477-82.

[18] Duchesne S, Pruessner JC, Collins DL. Appearance-based segmentation of medial temporal lobe structures. NeuroImage; 2002,17:515-532

[19] Thirion JP. Image Matching as a diffusion process: An analogy with Maxwell's demons. Medical Image Analysis, 1998; Vol 2 (3), 243-260

[20] Kim SH, Lee JM, Kim HP, Jang DP, Shin YW, Ha TH, Kim JJ, Kim IY, Kwon JS, Kim SI. Asymmetry analysis of deformable hippocampal model using the principal component in schizophrenia. Hum Brain Mapp. 2005 Aug; 25(4):361-9

[21] Shenton ME, Gerig G, McCarley RW, Szekely G, Kikinis R. Amygdalahippocampal shape difference in schizophrenia: the application of 3D shape models to volumetric MR data. Pyschiatry res. 2002 Aug 20; 115(1-2):15-35

Address for correspondence

Hussain Tameem
UCLA Medical Imaging Informatics
924 Westwood Blvd, Suite 420
Los Angeles, CA 90024 USA
Email: htameem@mii.ucla.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Automatic Image Modality Based Classification and Annotation to Improve Medical Image Retrieval

Jayashree Kalpathy-Cramer[a], William Hersh[a]

[a] *Department of Medical Informatics and Clinical Epidemiology, Oregon Health and Science University, Portland, Oregon, USA*

Abstract

Medical image retrieval can play an important role for diagnostic and teaching purposes in medicine. Image modality is an important visual characteristic that can be used to improve retrieval performance. Many test and on-line collections do not contain information about the image modality. We have created an automatic image classifier for both grey-scale and colour medical images. We evaluated the performance of the two modality classifiers, one for grey-scale images and the other for colour images on the CISMeF and the ImageCLEFmed 2006 databases. Both classifiers were created using a neural network architecture for learning. Low level colour and texture based feature vectors were extracted to train the network. Both classifiers achieved an accuracy of > 95% on the test collections that they were tested on. We also evaluated the performance of these classifiers on a selection of queries from the ImageCLEFmed 2006. The precision of the results was improved by using the modality classifier to re-sort the results of a textual query.

Keywords:

medical imaging, neural networks, image annotation, content-based image retrieval

Introduction

Medical images form a vital component of a patient's health record. Effective medical image retrieval systems can play an important role in aiding in diagnosis and treatment; they can also be effective in the education domain for healthcare students, instructors and patients alike. As a result of advances in digital imaging technologies, there has been a large growth in the number of digital images stored in recent years. In addition to the Picture Archival and Communication Systems (PACS) that are becoming omnipresent in hospital and clinics, there are numerous on-line collections of medical images. On-line atlases of images can be found for many medical domains including dermatology, radiology and gastroenterology. The sheer volume of medical image data provides for numerous challenges and opportunities in the arena of medical image retrieval.

Historically, the task of indexing and cataloging these collections has been performed manually. This is an arduous and painstaking task, and is prone to errors. Consequently, there is a desire to be able to automate the task of indexing these collections with a goal to improve the ability to search and retrieve relevant documents.

Medical image retrieval systems have traditionally been text- based, relying on the annotation or captions associated with the images as the input to the retrieval system. The last few decades have offered advancements in the area of content-based image retrieval (CBIR) [1]. CBIR systems have had some success in fairly constrained medical domains, including pathology, head MRIs, lung CTs, and mammograms [2]. However, purely content-based image retrieval systems currently have limitations in more general medical image retrieval situations, especially when the query includes information about pathology [3, 4]. Mixed systems (using both textual and visual techniques) have demonstrated improved retrieval performance, especially with regards to precision at the top of the list [4].

Medical image databases used for image retrieval or for teaching purposes often contain images of many different modalities, taken under varied conditions with variable accuracy of annotation. This can be true for images found in various on-line resources, including those that access the on-line content of journals[1].

Image modality is an important, fundamental visual characteristic of an image that can be used to aid in the retrieval process. However, the annotations or captions associated with images often do not capture information about the modality. Images that may have had modality associated with them as part of the DICOM header can lose that information when the image is compressed to become a part of a teaching or on-line collection. There have also been reported errors in the accuracy of DICOM headings [5].

The medical image retrieval task within ImageCLEF has provided both a forum as well as test collections to benchmark image retrieval techniques. The ImageCLEF campaign has been a part of the Cross Language Evaluation Forum since 2003 [3]. CLEF itself is an offshoot from the Text REtrieval Conference (TREC). In 2004, ImageCLEFmed, a domain–specific task, was added to evaluate medical image retrieval algorithms and techniques.

1 http://gm.arrs.org/ (accessed 3/26/2007)

Approaches combining both visual and textual techniques for retrieval have shown some promise at medical image retrieval tasks [3]. In 2005, a medical image annotation task was added to ImageCLEF. The goal of this task was to correctly classify 1000 test images into 116 classes given a set of 10,000 training images. The classes differed primarily in anatomy and view of the image. It should be noted, however, that these images were primarily of a single modality (X-rays).The goal of the ImageCLEF medical image retrieval task of 2006 was to retrieve relevant images for thirty topics from a test collection of about 50,000 annotated images of different modalities. These tasks were divided by the organizers into those expected to be amenable to textual, visual, or mixed retrieval techniques.

We participated in both the medical image retrieval and the automatic medical image annotation tasks at ImageCLEF 2006 [6, 7]. The techniques developed for those tasks have been extended for the more general task of medical image modality classification and annotation.

Using medical image modality for image annotation and retrieval has recently been studied. Florea et al [8] have compared the efficacy of two different systems (MedIC and MedGIFT) in classifying the modality of a database with six standard modalities for radiology and nuclear medicine images.

In this paper, we compare the results obtained on our system with those described in previous publications [8] for the six modalities of the CISMeF database. We will also extend this technique to classify colour images from the ImageCLEF medical retrieval task collection [6] into six categories. We will finally report on the improvement in precision that we observed for a selected number of tasks of the ImageCLEF medical retrieval task for 2006 by incorporating the modality classifier in series with a text-based retrieval system.

Methods

We employed a supervised machine learning approach to problem of medical image modality classification using a hierarchical classification scheme as seen in figure 1. There were two primary databases that were used to create and test the classifiers. We worked with a small subset of the CISMeF database as the primary target for our grey-scale (radiographic and nuclear medicine) image classifier [9]. This database had a set of 1332 images classified into one of six classes based on modality. These include angiography, computerized tomography scans (CT), X-ray, Magnetic resonance (MRI), ultrasound, and scintigraphy. The images in this database had been acquired under differing conditions over a long period of time. Consequently, there was considerable intra-class variation in quality, size, contrast, illumination and background.

The imageCLEFmed database contains 50,000 images of differing modalities, including radiography and nuclear medicine, as well as microscopic and histopathological images, photographs and gross pathology images, power point slides, electroencephalograhical images (EEGs) and electrocardiograms (ECGs), as well as a few miscellaneous images.

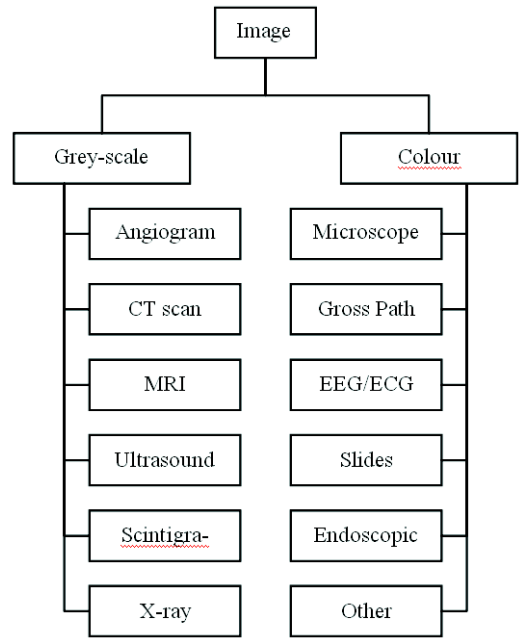


Figure 1 - Hierarchical classification scheme for images

A neural network-based scheme using a variety of low level, primarily global image features was used to create a six-class classification system for the grayscale images. The multilayer perceptron architecture used a hidden layer of approximately 50-150 nodes. The classification system was created in MATLAB[2], in part using several routines modified from the Netlab toolbox[3].

We experimented with a variety of feature vectors as inputs to the network. A combination of texture and intensity histogram features provided the best classification [10, 11]. All images were first resized while maintaining the aspect ratio such that the smaller dimension was 256 pixels. The image was divided into five overlapping blocks. Grey level correlation matrices were computed for each block using four angles and an offset of 1 pixel. Contrast, correlation, energy, homogeneity and entropy were calculated for each matrix. A quantized grey scale histogram was then appended resulting in a 132-dimension feature vector for each image for the texture. All inputs to the neural network (the image feature vectors) were normalized using the training set to have a mean of zero and variance of 1.

The 1332 images in the database were randomly split into a training set of 1000 images and a test set of 332 images. A small random subset of the training images was initially used to create the classifier (200 images). The classifier

2 www.mathworks.com (accessed 3/26/2007)
3 http://www.ncrg.aston.ac.uk/netlab/index.php (accessed 3/26/2007)

was then applied to the entire training set and images that were misclassified were then added to the images used to refine the classifier. The classifier was finally tested on the test images.

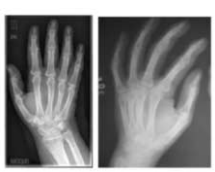
A similar scheme was used to create the classifier for colour images. We believe this novel idea can improve the retrieval performance of purely textual systems or for images for which the associated modalities are not known. Although modality detectors for grey-scale medical images have been reported [9], we are unaware of a similar effort for classification of other categories of medical images like those produced by histo-pathology and endoscopy. The images used for this classification task were taken from the test collection of images used in the ImageCLEFmed retrieval task. 2250 colour images in this collection were broadly categorized into six categories as microscopic, gross pathology, EEG/ECG or other charts, powerpoint slides, endoscopic images and other. There was considerable intra-class variability in this dataset. These 2250 images were again split randomly into training (1750) and test images (500). A similar training methodology to that described above was used to incrementally improve the classifier, starting with a smaller subset of the training database.

A two-layer architecture with 25-150 hidden nodes was used for the neural network. The feature vector in this case consisted of colour histogram features, as well as texture features obtained using the grey level correlation matrices. The image was split into 9 uneven blocks. Colour histogram properties of image after conversion into the L*A*B* colour space were calculated, while texture features were calculated after converting the image to grey-scale

These neural network classifiers can be created to further classify images within a given modality. For instance, x-ray images could now be classified to account for anatomy. Anatomical classifiers were used in the automatic annotation task at ImageCLEFmed.

The tasks had been stated in English, German and French, and had and provided example images. All but three of the tasks stated the desired modality of the image to be retrieved. Two examples of the tasks are shown in figure 2.

Show me images of a hand x-ray.
Zeige mir Röntgenbilder einer Hand.
Montre-moi des radiographies de la main.



Show me blood smears that include polymorphonuclear neutrophils. Zeige mir Blutabstriche mit polymorphonuklearer Neutrophils. Montre-moi des échantillons de sang incluant des neutrophiles polymorphonucléaires.

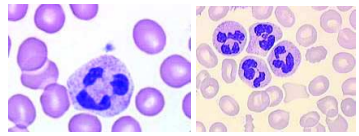


Figure 2 - Sample textual and visual queries at ImageCLEFmed 2006

Once our classifiers had been trained to achieve >95% classification accuracy, they were tested on a random subset of the ImageCLEFmed topics.

The schematic of our modified retrieval system is shown below. The query was initially fed to our Lucene[4] based text retrieval system. The queries were manually edited by one of the authors. The resulting images were subsequently classified by the hierarchical classifier for modality. Images of the desired modality (as stated in the query or as discerned by the automatic classifier based on the sample images) were moved to the top of the list while maintaining the ranking of the textual system within a class.

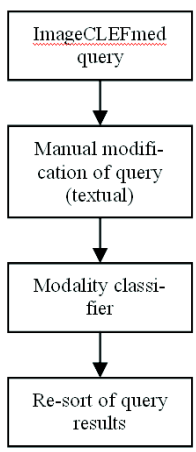


Figure 3 - Image retrieval system used for the ImageCLEFmed 2006 test collection

We compared the results of our purely textual system with that including the addition of the modality classifier.

Results

A classification accuracy of 96.4% was achieved on the CISMeF database. The confusion matrix suggests that the primary misclassification occur between the MRI and CT scan classes. This is not surprising as these classes are visually quite similar. Florea et al [8] have reported similar results both in terms of accuracy and inter-class misclassification patterns. The classification of grey-scale medical images into commonly occurring modalities using low level image features and machine learning techniques appears to be a tractable task. We expect to achieve over 98% accuracy with further refinement of our machine

4 http://lucene.apache.org/ (accessed 3/26/2007)

learning approach by the use of more advanced cross-validation, bootstrapping, boosting or bagging techniques.

Preliminary testing of the classifiers on 2250 colour images of the imageCLEFmed test collection resulted in a modality classification accuracy of 98.6%. Most of the misclassifications involved the "other" class with contained a set of miscellaneous images not belonging to the other five specific categories

The colour modality classifier was tested on a small random subset of the ImageCLEFmed 2006 topics. The topics for imageCLEFmed 2006 fell into three categories (visual, mixed, semantic) consisting of 10 tasks each. Although visual techniques had, in general, performed extremely poorly at the semantic tasks, use of some visual information (primarily modality) was shown to increase the precision [4].

Analysis of our textual results indicated that in many queries, especially those of a visual or mixed nature, up to 75% of the top 1000 results were not of the correct modality. A compelling example is given in figure 4 and table 1. Only 90 of the top 2000 images returned by the textual query were of the desired modality.

Task 1 - Show me images of the oral cavity including teeth and gum tissue

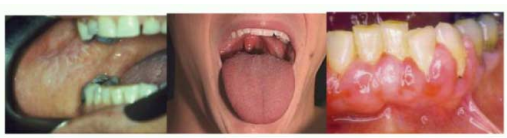


Image type	Number of images
Total returned by textual query	2000
Grey-scale	1831
Photograph/gross pathology	90
Microscope	71
Other	8

Figure 4- Sample query suitable for visual retrieval at ImageCLEFmed 2006

These images were then classified using our modality classifier. The ranked list of retrieved images was resorted taking into account the desired modality based on the query.

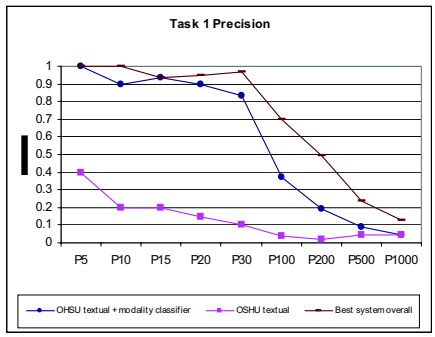


Figure 5 - Improvement in precision resulting from modality classification.

Figure 5 plots the precision for varying number of documents retrieved for the purely textual system, the improvement with the use of the modality classifier and the overall best system (mixed visual and textual based) that participated in ImageCLEFmed 2006. This increased the precision of the query as seen in figure 5. The improvement in precision at the top of the ranked list (P5 – P200) is better with the use of the modality detector compared to a purely textual search. We should note that a perfect modality classifier will only improve the precision of the search and not the recall if it is applied in the serial manner described above. The mean average precision (MAP) would still be limited by the number of relevant images that are retrieved by the textual search (recall of the textual search).

Even in searches that are expected to be semantic, we see an improvement in precision by using the modality classifier as seen in figure 6 and 7.

Task 2 - Show me microscopic images of tissue from the cerebellum (semantic query)

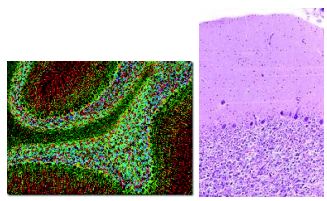


Image type	Number of images
Total returned by textual query	2000
Greyscale	1476
Photograph/gross pathology	408
Microscope	116

Figure 6 - Sample query suitable for visual retrieval at ImageCLEFmed 2006

The precision of this search was similarly improved by the use of the modality detector as seen in figure 7.

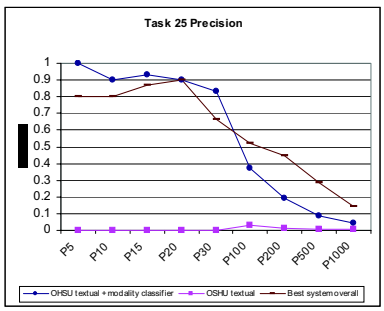


Figure 7 - Improvement in precision resulting from modality classification

Four of the six tasks tested showed improvement in precision by the use of the modality detector for colour images. There were two tasks amenable to textual methods for which there very little change in precision with the addition of the modality information.

We plan on testing the performance of the modality detector on the complete set of tasks for ImageCLEFmed 2005 and 2006. We also intend to index the entire collection of 50,000 images used in the ImageCLEFmed test collection using the modality classifier. Information about the class membership of an image will be added to the metadata. This should improve the performance of the retrieval in two ways. Clustering of the data by modality and perhaps anatomy will speed up the search process as fewer documents will have to be compared to the query image/text. Secondly, we expect that the overall precision of the search will improve by considering the modality of the image that is desired by the user. However, we can expect a small degradation in the recall due to potentially misclassified images not being searched.

Conclusion

We have developed a neural network based, hierarchical classifier for the modality classification of medical images. This system can classify colour images including histo-pathological and endoscopic images, and photographs as well as grey-scale (radiological and nuclear medicine). The classifier uses a histogram and texture properties as inputs to the two level neural network. This classifier results in a classification accuracy of greater than 95% for the grey-scale images of the CISMeF database as well as a selection of colour and grey-scale images from the ImageCLEFmed database. The use of this classifier increases the precision of retrieval of our primarily text based retrieval system by moving images of the desired modality to the top of the ranked list.

Acknowledgments

The authors would like to thank Henning Müller for his valuable input in discussions as well as providing access to the CISMeF images. We would also like to thank Jean-Nicholas Dacher at the Rouen Hospital, France, for proving the CISMeF images used in this paper. We would like to thank Dr. T. Lehmann at RWTH Aachen University, Germany, for access to the IRMA image database. We acknowledge the support of NLM Training grant 5T15 LM07088-15 and NSF Grant ITR-0325160.

References

[1] Smeulders AWM, Worring M, Santini S, Gupta A, Jain R. Content-based image retrieval at the end of the early years. IEEE Transactions on Pattern Analysis and Machine Intelligence 2000;22(12), 1349-80.

[2] Müller H, Michoux N, Bandon D, Geissbuhler A. A review of content-based image retrieval systems in medicine – clinical benefits and future directions. International Journal of Medical Informatics 2004; **73**, 1-23.

[3] Müller H, Deselaers T, Lehmann T, Clough P, Hersh W. Overview of the ImageCLEFmed 2006 medical retrieval annotation tasks, Evaluation of Multilingual and Multi-modal Information Retrieval, Seventh Workshop of the Cross-Language Evaluation Forum, CLEF 2006, editors: Peters C, Clough P, Gey F, Karlgren J, Magnini B, Oard DW, de Rijke M, Stempfhuber M, LNCS 2006,Alicante, Spain, to appear.

[4] Hersh W, Kalpathy-Cramer J, Jensen, J. Medical Image Retrieval and Automated Annotation: OHSU at ImageCLEF 2006. Working Notes for the CLEF 2006 Workshop, Alicante, Spain. http://www.clef-campaign.org/2006/

[5] Güld MO, Kohnen M, Keysers D, Schubert H, Wein BB, Bredno J, Lehmann TM, Quality of DICOM header information for image categorization, Proceedings SPIE, 4685, 280-287, 2002.

[6] Hersh W, Müller H, Jensen J, Yang J, Gorman P, Ruch P. Advancing biomedical image retrieval: development and analysis of a test collection. J Amer Med Inform Assoc 2006; 13(5).

[7] Lehmann TM, Güld MO, Thies C, Fischer B, Keysers D, Kohnen M, Schubert H, Wein BB. Content-based image retrieval in medical applications for picture archiving and communication systems, in Medical Imaging, SPIE Proceedings, 5033, 440-451, San Diego, California, 2003.

[8] Florea F, Müller H, Rogozan A, Geissbühler A, Darmoni S, Medical image categorization with MedIC andMedGIFT. Medical Informatics Europe (MIE 2006).

[9] Douyere M, Soualmia LF, Neveol A, Rogozan A, Dahamna B, Leroy JP, Thirion B, Darmoni SJ. Enhancing the MeSH thesaurus to retrieve French online health resources in a quality-controlled gateway. Health Info Libr J 2004; 21(4):253-61.

[10] Haralick R. Statistical and structural approaches to texture. Proceedings of the IEEE 67**,** 786-804, 1979.

[11] Howarth P, Ruger S. Evaluation of texture features for content-based imageretrieval. In: Proceedings of the International Conference on Image and Video Retrieval, Springer-Verlag, pp. 326-324, 2004.

Address for correspondence

Jayashree Kalpathy-Cramer
5th Floor, Biomedical Information Communication Center
3181 S.W. Sam Jackson Park Rd.
Portland, Oregon 97239-3098

Email: kalpathy@ohsu.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Quantification of Myocardial Perfusion for CAD Diagnosis

Hartmut Dickhaus[a], Markus Erbacher[a], Helmut Kücherer[b]

[a] *Department of Med. Informatics, University of Heidelberg, Heidelberg, Germany*
[b] *Department of Cardiology, University Hospital, Heidelberg, Germany*

Abstract

We introduce a computer based algorithm for objective quantification of myocardial perfusion to support the diagnosis of cad patients. This new method is based on conventional cine angiographic films. In order to achieve maximal quality of the digital subtraction angiography images, the sequence is synchronized with the ECG. Optionally, the digital images can be motion compensated by a two step matching method. The spatio-temporal spread of blood, or the so-called blush, through the microvasculature to the myocardium - indicated by dye injection – represents a characteristic pattern for the myocardial perfusion. This dynamic temporal pattern is characterized by typical features as the maximal value of blush intensity, of increase and of decrease velocity which correspond with the different phases of flooding in and washout. On the basis of 100 different temporal blush profiles, an algorithm is established which classifies the acquired blush patterns into 4 different grades.

Keywords:

angiography, blush grade, CAD diagnosis, myocardial perfusion

Introduction

Currently, the coronary angiography is still the gold standard for coronary artery disease (CAD) diagnosis, although other methods have been proposed and discussed [1]. In almost all cases, the required invasive procedure of inserting a catheter for dye injection is combined with a percutaneous transluminal coronary angioplasty (PTCA). Therefore, the angiographic procedure allows an immediate estimation of the success of the therapeutic intervention.

For quantifying complete and sustained reperfusion of the infarcted myocardium and prognostic statements as for identifying patients at high risk, myocardial blood flow, expressed in so-called blush grades, is much more appropriate than any other angiography related measure. This is demonstrated by Stone et al. [2].

However, so far only qualitative descriptions for different extends of blush or myocardial reperfusion exist. Gibson et al. [3] presented 4 different grades which classify the perfusion in relation to its temporal dynamic and intensity. For example: grade 1 is defined as: "dye slowly enters but fails to exit the microvasculature. There is no blush or opacification on the myocardium in the distribution of the culprit lesion that fails to clear from the microvasculature and dye staining is present on the next injection 30 sec later."

Despite the principal advantages of blush characterization, qualitative descriptions like these are difficult to apply in an objective and reproducible stratification. For that reason, we tried to establish a computer assisted procedure to quantify blush grades corresponding with those of Gibson, however, in an objective and formal description. Furthermore, we developed a computer aided tool to visualize the spatial and temporal spread of the adelomorphic blush, of the myocardium from cineangiographic films after dye injection into the arteries. Under the control of the cardiologist, a specific blush grade is assigned for a specific myocardial region related to the three main supplying coronary arteries.

Materials and methods

In this study, 100 films from patients with various extent of CAD recorded in different projections, are quantitatively examined before and after PTCA. Because the heart is supplied by three different arteries: right coronary artery (RCA), left artery descending (LAD) and left circumflex (LCX), we look for blush occurrence in correspondent myocardial areas supplied from these vessels.

In order to enhance the contrast between blush and the arteries and surrounding tissue, images after dye injection have to be subtracted from that prior to injection, respectively in standard position. The initial image without dye is called a mask and is mostly established by averaging two or three consecutive frames in order to smooth small artifacts and noise. The difference of images shows the highest contrast between filled vessels, microvasculature and the surrounding, if we take the logarithmically transformed angiograms. In order to consider the motion of the heart and its vessels as well as the inserted catheter during the heart cycle, one has to choose images for subtraction which correspond to identical heart geometry and position. This can be achieved by synchronization of the image acquisition to the R peak of the simultaneously recorded ECG. Fortunately, the time interval between two cycles,

about 1 sec, is relatively small, so that the resulting time course of dye spreading is still sufficiently documented.

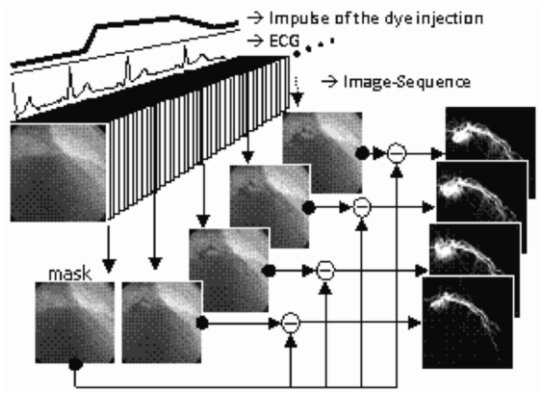


Figure 1 - Acquisition of ECG synchronized frames from the cine angiographic film. The black marked frames related to the R peak of every cardiac cycle are used for creating the sequence of difference images, displaying the spreading blush across the heart.

In figure 1 the acquisition procedure is schematically demonstrated. With increasing time we get a sequence of difference images displaying the temporal progression of dye and its spatial spreading through the arteries and small vessels into the microvasculature and the myocardium. The quality of these images demonstrating the blush is highly dependent to any change of body position or motion of inner organs, like the diaphragm, during breathing. For this reason, the patient is asked to stop breathing for about 20 sec, and the camera as well as the operating table have to be fixed for that period. However, in clinical routine it is difficult to accomplish these conditions completely. Therefore, we apply additional compensation procedures in two consecutive steps.

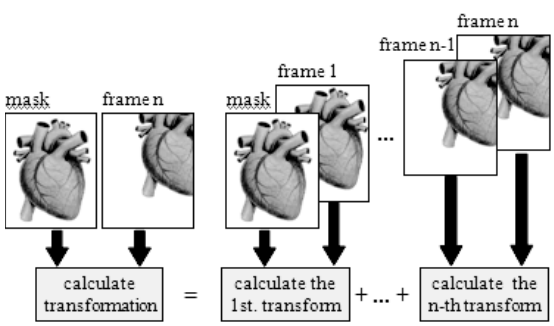


Figure 2 - Schema of the registration procedure between frame n and the mask: the registration is split into n single consecutive transformations and finally the sum of all transformations is performed.

Movements of the operation table or of the camera adjusted by the cardiologist for optimal viewing conditions are primarily compensated by a rigid matching procedure. Secondly motion of the inner organs particular caused by breathing needs an elastic compensation. In most cases the table or camera movements are larger but can be corrected more easily than the breathing effects. This two step procedure minimizes the extent and time for the elastic registration [4].

One general problem characterizes both registration steps, even the continuous spread of dye during the digital acquisition procedure. This means that besides the mentioned motion artefacts we have to consider differences between succeeding frames resulting from the dye distribution. In order to minimize this problem we apply the matching procedures always between two consecutive image frames. That means that the registration is performed every time relative to the previous frame. Finally all single transformations are summed up. Figure 2 demonstrates this process schematically. To calculate each transformation step, the rigid one as well as the elastic one, primarily the vessels have to be segmented for each frame. The pixels identified as vessels are then skipped during the calculation, in order to minimize the problems caused by the spreading dye.

The spatial distribution of the blush is represented by the gray value averaged from myocardium pixels excluding the arteries. That means we have to identify all myocardium pixels belonging to the blush. This task is very hard even for an experienced cardiologist. We perform this segmentation by an interactive procedure. For all pixels, we trace the intensity profile, respectively the gray value, over time. The incoming wave of dye after injection follows primarily the main coronary arteries, flows through the smaller vessels into the microvasculature, and reaches the myocardium. The washout is collected in the venous vessels. This time dependent process can be represented in its spatial distribution by detecting the maximal values of the intensity profiles of all image pixels. All pixels identified by their maximal gray values at a certain time instant represent more or less a specific structure of the pathway to the venous system. As it is demonstrated in Figure 3, it is easy to choose a color map representing the arteries map c) or d) or the perfused myocardium, pixel map f). For all segmented pixels of map f) we calculate the average gray value, cycle for cycle and plot the corresponding time course of blush intensity.

The cardiologist has to position a specific ROI at a region of the myocardium besides the arteries where a blush would be expected under normal conditions. We also calculate for this specific region the average gray values for different instants of time as characteristic time courses for a restricted perfusion. The different time dependent profiles have to be characterized by a few parameters which can be used to distinguish between different grades of blush. Furthermore, we would like to define parameters which can be easily interpreted in context with the existing qualitative characterizations of Gibson, and finally they should be related to the pathophysiological background.

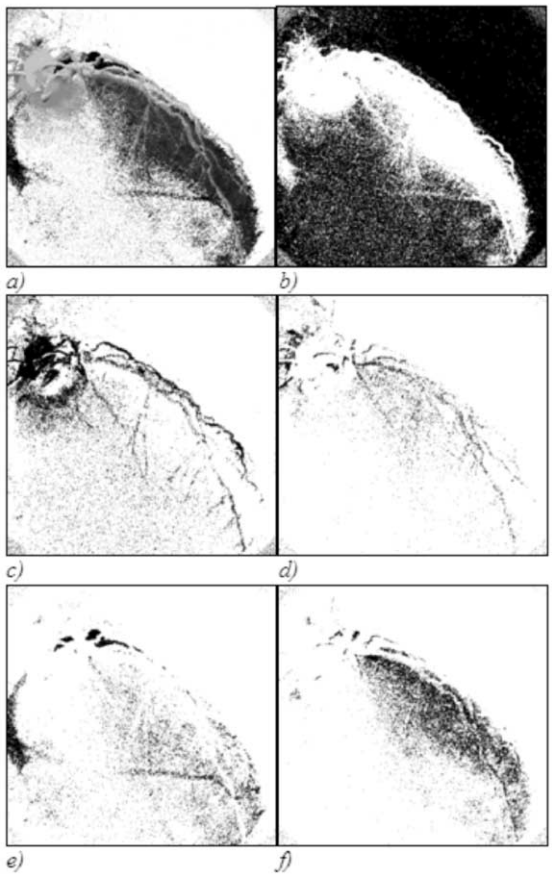


Figure 3 - Color maps for characterizing of specific structures during blood perfusion. The temporal incidence of the maximal dye intensity is mapped on to the pixel matrix. For different instances of time, corresponding with map b) to f), we display different anatomic structures which are perfused at this time; a) shows all maps superimposed: background, artery (LAD), branching, myocardium, vein; b) - f) the separated structures. Map f) is of particular interest because it represents the pixels of the myocardium.

We defined 4 different parameters for each intensity profile which are displayed in Figure 4: the maximal intensity Gmax; the time of Gmax; the maximal slope Imax of the rising profile, which corresponds with the maximal rising velocity of the perfusion; and the maximal slope Dmax of the falling profile, which corresponds with the maximal outwash velocity. For all calculated blush profiles, these four parameters are extracted as a typical characterizing feature set.

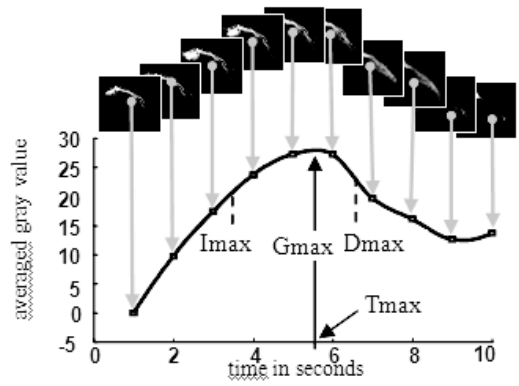


Figure 4 - The time course of averaged gray values from corresponding areas of the perfused myocardium. 4 specific parameters: Imax: Increase-slope, Dmax: Decrease-slope, Gmax: maximal gray-value, Tmax: time of Gmax are indicated

Results

For all 100 films the corresponding time courses are established and automatically characterized by Gmax, Tmax, Dmax and Imax. It is of particular interest to demonstrate the difference of blush spreading before and after a PTCA intervention by different temporal patterns.

Different degrees of stenosis are related with different extensions of blush and different patterns of dynamic profiles. The patient of Figure 5 had an occlusion of RCA. After PTCA the artery looks absolutely normal (Figure 5, lower trace). By positioning of a ROI in the supply area of this vessel we calculate the corresponding blush profile with its characterizing parameters. However, for the same region before PTCA we found a very flat profile (Figure 5 upper trace), which shows us that there is not any perfusion of the myocardium. Nevertheless, the reperfusion after PTCA is obvious. In this case we have consistent results with the re-opened artery and the blush profile. However, sometimes the situation demonstrated on the angiograms is not so clear. For these cases we get much more information by the quantified blush profiles.

In order to relate the various blush profiles to blush grades a non-parametric classification procedure based on the four specific parameters is under development. First promising results are in good agreement with the subjective grading of clinical experts.

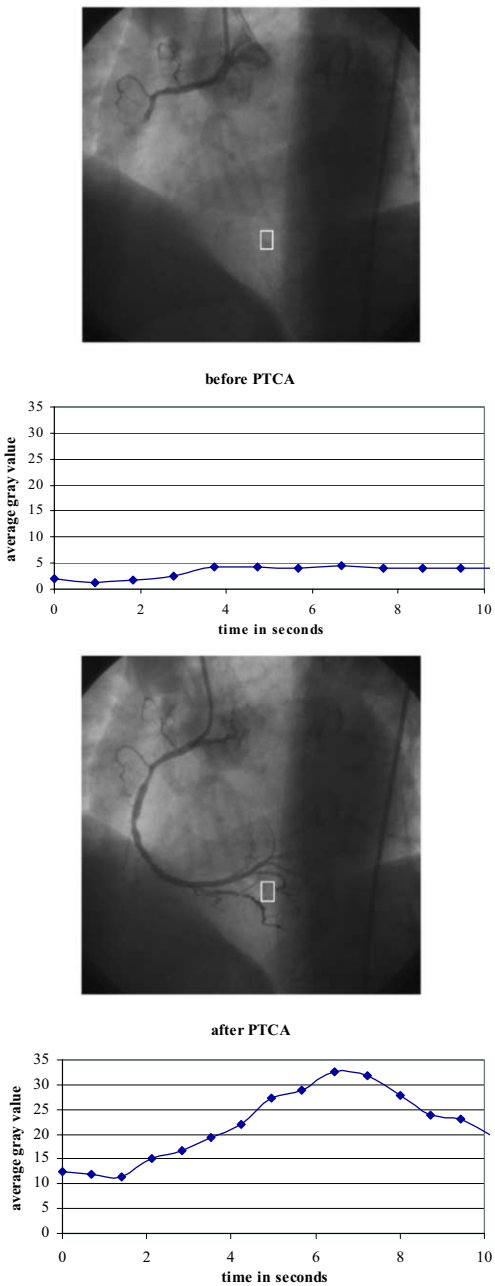


Figure 5 - Angiograms before and after PTCA of a patient with a complete stenosis of 100% for RCA. The corresponding temporal blush profiles document obviously the improvement of myocardial reperfusion

From the 100 digitized angiographic film sequences we could evaluate only 60 without considerable motion compensation. The introduced two step registration method allowed an additional evaluation of 30 more series. The remaining 10 films show such considerable motion artefacts of the diaphragm that an evaluation was not possible. In all of these cases the diaphragm moved over the whole hearth and modified its grayvalues significantly. Figure 6 shows the effect of motion compensation.

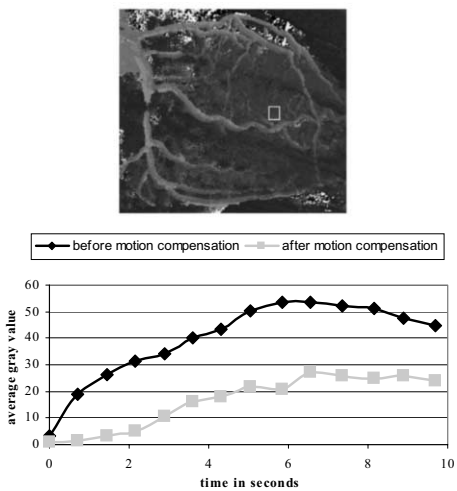


Figure 6 - The left image shows the extent of usually recorded motion artefacts whereas the right image demonstrates the benefit of motion compensation. The time courses before and after motion compensation are plotted in the same diagram below

Discussion

Our presented approach on the basis of digital subtracted angiographic images is an interactive procedure which helps the cardiologist to assign blush grades in strong relation to what he is used to do. He gets an intuitive impression of the spreading blood after the dye injection by displaying the processed colormaps. It is easy to confirm a quantitative hypothesis by looking precisely to selected areas of interest of the myocardium, and its temporal gray value profile or to integrate the average dye activity of the whole area.

The advantages of the developed system are given in a highly adaptable degree of automation. Dependent on the quality of angiograms, the more or less experienced user can interfere in different kinds. For example, simple adjustment of the contrast or corrections of motion artefacts lead to an improved quality of the digitally subtracted image sequence. The automated classification procedure is also easily adaptable to other types of blush patterns. Up to now it was our aim to implement the descriptions for blush grading of experienced cardiologists as it was proposed by Gibson. However, if we wish to characterize any other specifity of the perfusion, we can as well introduce more quantitative features related to the blush patterns, for example temporal measures. Furthermore, other approaches of motion compensation could be incorporated and evaluated for their practicability in respect of typical

angiographic artefacts. Of course, we have to look for more sequences to incorporate them in our learning sample. Therefore, we developed an intuitive user interface which helps the cardiologist to follow the different steps of the acquisition procedure and which offers him the various facilities to derive a reliable diagnostic statement. During the current evaluation of our system we are discussing the usefulness of a more automated selection of the ROI for blush profile quantification. However, up to now the experienced cardiologists are not convinced that this feature increases the clinical benefit significantly.

References

[1] Bibra H. Assessment of myocardial perfusion - role of contrast echocardiography. Z. Kardiol 2001: 90: pp. 848-859

[2] Stone GW, Lansky AJ, Mehran R, et al. Beyond TIMI 3 Flow: The Importance of Restored Myocardial Perfusion for Survival in high risk Patients undergoing primary or rescue PTCA. J Am Coll Cardiol 2000: 35: 403A

[3] Gibson CM, et al. Relationship of TIMI Myocardial Perfusion Grade to Mortality after Administration of Thrombolytic Drugs. Circulation, 2000: 101(2): pp. 125-130

[4] Modersitzki J. Numerical Methods for Image Registration. Oxford University Press, 2004

Address for correspondence

Hartmut.Dickhaus@med.uni-heidelberg.de
Prof. Dr. H. Dickhaus
Institute for Medical Biometry and Informatics
Department of Medical Informatics
University of Heidelberg
Im Neuenheimer Feld 400
D-69120 Heidelberg Germany
Tel.: 0049 6221 567483
Fax: 0049 6221 564997

This page intentionally left blank

Chapter 11.

Education and Training

This page intentionally left blank

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Enabling the Safe and Effective Implementation of Health Informatics Systems – Validating and Rolling Out the ECDL/ICDL Health Supplement

Michael J. Rigby[a], Carol Hulm[b], Don Detmer[c], Luca Buccoliero[d]

[a] *Centre for Health Planning and Management, Keele University, United Kingdom*
[b] *British Computer Society, Swindon, United Kingdom*
[c] *American Medical Informatics Association, Bethesda, MD, United States of America*
[d] *SDA Bocconi and CeRGAS Centre for Research on Health Care Management, Bocconi University, Milan, Italy*

Abstract

Sound understanding by end-users (health professionals and support staff) of key e-health principles and challenges is essential for the safe, effective, and sustainable use of health informatics systems. This is separate from, and ideally a precursor to, training on a specific system. However, hitherto this aspect has been little understood or addressed. Over the last few years, the concept of a customised Health Supplement to the well-established European/International Computer Driving Licence has progressed from idea to reality, through initial support for consultation by the then UK NHS Information Authority, followed by establishment of an international Expert Group by the global ECDL/ICDL Foundation. As a result, the ECDL/ICDL Foundation has developed a formal syllabus. This paper reports successful piloting, and progress in the development of local training and testing materials and national implementation plans, in three countries on two continents.

Keywords:

health informatics education; end-user; ECDL; ICDL; health informatics sustainability

Introduction

As defined by the theme for Medinfo 2007, there is rapidly growing recognition and dependence in all countries on the future major role of health informatics systems in enabling national health systems to be increasingly effective, efficient, and evidence based. Whilst each country starts from a different base line, all are moving forward at their local pace in increasing the use of informatics systems.

Rightly, the priority is increasingly for sustainability, which must include the reliable and efficient use of such systems, in line with intention and with capability. However, the fundamental oversight is to focus solely on systems design and implementation, important though they are. What is so often overlooked is the fact that systems are only as good as their users' understanding and capacity to use them appropriately.

Health informatics systems are a vehicle for collecting, processing, and making available relevant health information at the point of need. Therefore such support systems should not dominate or hamper health practitioners' activities except so far as they can make them more effective or efficient. However, all too often end users – health practitioners and their support staff – can find such systems threatening and restricting, and in turn this means that systems are used inefficiently and ineffectively, and either perform badly or fall into disuse.

Resistant mature staff

Though often a surprise to many, this picture of end-user poor compliance is well documented over a period of time, and thus should be anticipated [1]. Health professionals have in round terms a professional lifespan of approximately forty years. In nearly every country those practicing for more than ten years will have undertaken their basic professional education and early practice in an environment based on paper-based records. In many countries of the world electronic systems will only have been brought in even more recently. Furthermore, the most senior and experienced health practitioners – usually the professional champions and leaders – will certainly have been educated ahead of the e-health revolution. It is into this workforce environment that governments, health policy makers, and health informatics system advocates are seeking to introduce radical and comprehensive electronic systems.

An atypical IT system user population

A further complication is that not only are health information system end users more well versed in paper-based systems, in many other ways too they differ from the end users of almost all other organisation-wide modern electronics records systems, such as those in banking or the travel industry [2,3]. Not only are health users senior and mature staff rather than new entrants; they are totally dependent on information systems in order to carry out their principal daily business. However, this core work is not itself focussed on the information system - in general the use of the information system only forms between 10% and 20% of their duties - and to them this is a low-skill element of their work. Thus senior doctors, nurses, and other

Selected for best paper award.

health professionals will not give the same level of importance to being trained in new-generation information systems as they will to new clinical or healthcare techniques. Yet, even though being only part-time users, they are senior and often autonomous employees, and they have professional accountability over and above their duty to their immediate employer. Thus they will only be effective, reliable and regular users if they both understand the way to harness a system, and trust in its integrity and that of all partner users (who are largely unknown to them individually). The deeper analysis undertaken of these differences of health informatics system end user profiles compared to others system user profiles has been published [2, 3].

Change-inducing systems

Added to these challenges, health informatics systems in themselves do not support clinical practice in the simplistic ways that telephones support and replace meetings and written correspondence, or dictation systems replace the need for the physical presence of a secretary. Instead, health information systems require a very different pattern of working, ranging from new data recording processes, through to how to search a past record (which is navigated very differently if it is held in computer files compared with if it is a large collection of paper charts in a physical folder).

Informatics system sustainability but practice challenge

Therefore sustainability is highly dependent as much upon the pattern of use as it is on the pattern of design, yet is affected also by the user population's characteristics. Thus it is clearly naive to expect the complete cohort of the most senior and experienced practitioners in their country universally to welcome and endorse new systems which require radically different physical and cognitive skills, and which require immediate changes to patterns of data recording and assimilation honed over a life time. Thus, however good the system introduced, it will not be sustainable if the principal stakeholders are not comfortable with issues.

Counteracting the principal risks of health information systems

Apart from recognising the challenge to traditional practice of new health informatics systems, it is also important to recognise the inherent risks these bring if introduced without due preparation. Moreover, professionals will be aware in general terms of these risks, and be likely to militate against introduction of such systems with good intent unless they feel that these risks have been addressed and controlled. There are three types of such risk, as shown below.

Risks generated by the need for new skills

The use of a health informatics system requires a radical range of new skills. These commence with the basics of operating any computer system, through the skills required to record data electronically as apposed to by hand writing or filing a chart, to the skills needed to navigate a record which is stored in a highly structured and efficient way but which needs a new mind set in order to negotiate it effectively to find key and relevant information items. It would be unreasonable to expect a surgeon to use a new type of instrument or a radically new surgical procedure without adequate training, yet governments and policy makers worldwide are inappropriately labelling as "obstructive" health practitioners who are reluctant to change information management approaches with which they are familiar, for ones which they find unknown and intimidating. Further, 'smart' systems may make good evidence-based calculations, recommend particular treatment patterns or warn against particular prescribing intentions. These are safe provided the end user understands the rationale in both the clinical and computing logic inbuilt, but carry risks if the end user does not understand and know how to ascertain that logic.

The risks of new constraints

A key aspect of most health informatics systems is the fact that they require a standard approach to the description of histories, investigations, results, diagnoses, and interventions - in other words, the benefits of standard terminologies and taxonomies should reduce ambiguity and render observations and findings interpretable accurately by all. However, the converse of that is that an individual's well-developed means of indicating valuable subjective information such as uncertainty, provisional views, or feelings as apposed to hard evidence are rendered impossible. This may either exclude uncertain information, or result in it being recorded with a spurious impression of certainty. Users may know the clinical approach and specific skills of colleagues who hand-record narrative information, but this authorship and personalisation may be lost with electronically captured and stored data.

Risks of misuse

The very strength of health informatic systems – that they can search and present information from very large databases extremely quickly – is a potential risk as well as being a core purpose. Files can become increasingly comprehensive, and information can be obtained about many people. Careless use of such information can lead to inappropriate divulgence of confidential information, and without safeguards there are clearly risks of an ethical or malicious misuse. Thus all end users need to be educated to avoid these risks, and to adhere to robust corporate policies to control usage.

For all these reasons it is therefore important for end users to be adequately educated as to how to use systems soundly and effectively. This is a key part of the sustainability of systems. Evidence (or even suspicion) of misuse of systems, or of poor clinical decision making because of inability to use systems, will provide a rapid means of ensuring their demise. Such evidence or suspicions may come either through professional sources, press reports, or collective patient anxieties.

The move to a health end-user qualification

The need to recognise end users

The educational needs of end users are very different from those of technical health informatics staff. Over a decade ago a European Commission Concerted Action entitled EDUCTRA identified the informatics educational needs of health professionals as being different from those of IT staff [4]. This work suggested a range of necessary learning outcomes for each group – though for health professional users they focussed primarily on basic curricula for new entrants. Subsequently the International Medical Informatics Association (IMIA) addressed this topic, and produced recommendations of what should form the basis of health informatics education globally for each of the two staff communities. Both before and since these recommendations, the prime focus has been on informatics education, with development of many formal courses, and on the introduction of some informatics training into basic health professional education. Neither of these groups, however, form the general body of the practice community to whom new organisation-wide informatics systems are introduced or imposed.

The concept of an end user qualification

To overcome these anxieties and risks it seems self-evident that an appropriate end-user educational programme, and related qualification, for health informatic systems users should be developed, but this was not being addressed. Meanwhile, virtually all countries in the world have a qualification requirement for drivers of motor vehicles or pilots of aeroplanes, as such equipment is seen as extremely beneficial yet extremely risky if misused through ignorance or lack of skills. It is seen as a societal responsibility to provide a qualifications framework and regulation, and a citizen responsibility to ensure qualification before becoming a user.

This approach has already been taken with the more general use of computers, with the development of the European Computer Driving Licence (ECDL) to a standard international curriculum, as most of the issues of using computers safely and effectively are generic and universal. This has now developed into the global International Computer Driving License (ICDL), available in virtually all countries of the world [6].

It therefore seemed logical to develop a specific supplement or module for the ECDL/ICDL, given the risks and responsibilities inherent in using such systems. This concept was first promoted in 1999 in a European context [7]. Subsequently, the idea was developed at conceptual level in more detail, and support gradually developed [8]. Of significance for Medinfo 2007, details were sought by a principal Australian health informatics journal [9].

Practical steps to development

Following these moves towards the development a health supplement to the ECDL/ICDL as the best means of meeting this need, and thereby ensuring sustainable and safe implementation of health informatics systems through education, assurance, and empowerment of end users, many practical steps have been made towards achieving this reality.

In 2004 the NHS Information Authority, the then lead body in this field for the National Health Service in England, agreed to support two consultation workshops – one for key opinion leaders in health informatics from eight European countries, and one for a range of delegates from the National Health Service across the United Kingdom. As a result of the strong enthusiasm at both these meetings, the European Computer Driving Licence Foundation (the global regulatory and licensing body for the ECDL/ICDL) agreed to consider formalising the development process. The ECDL Board endorsed this, and in 2005 an Expert Group was set up comprising representatives of six European countries and of the United States of America. The resultant recommended syllabus was signed off by the Expert Group in early 2006.

The ECDL/ICDL health supplement content

The final ECDL/ICDL Health Supplement consists of a competencies framework defining knowledge and skills the candidate needs to possess in order to operate a health information system safely. It excludes generic issues covered in health professional training or staff induction (such as basic principles of confidentiality). Regarding computer recording, it focuses on those aspects which are different, or have different emphasis or importance, in health applications.

The core contents of the syllabus are copyright to the ECDL/ICDL Foundation, and comprise the following topics:

- Concepts
  - Health Information Systems
  - HIS Types
- Due Care
  - Confidentiality
  - Access Control
  - Security
- User Skills
  - Navigation
  - Decision Support
  - Output Reports
- Policy and Procedure

For each topic a number of defined knowledge areas or competencies are specified. The content is designed to accommodate specific national language and terminologies, organisations, and legal and professional frameworks. The normal pattern of assessment will be electronic, through a testing framework available on line or by other electronic means. It is based on the assumption that the candidate will already be competent in basic computer user skills.

International trials and validation

Since the specification phase, rapid progress has been made in significantly different countries, with very differ-

ent health systems, different languages, and also different terminology and nomenclature within the same language group.

United Kingdom

In the United Kingdom, the British Computer Society as national licensee for the ECDL, and with a strong relationship with the National Health Service, organised piloting of the syllabus utilising an interim training manual and testing framework in six very different sites. These encompassed very different localities, and different healthcare environments ranging from primary care through secondary care to mental health, and different health professions from research staff to medical consultants, and also health informatics and health data experts. This pilot involved 84 persons, who were all very positive on the value of the knowledge and competencies covered in the syllabus. The only significant comments received were about the interim testing framework, which was only ever intended to be temporary in order to facilitate consideration of the syllabus. More detailed reporting of these results is in press [10].

Consequent upon these successful pilots, a full electronic training resource has been developed from the interim one, and a definitive electronic testing framework built. The ECDL Health Supplement was launched to English NHS staff by NHS Connecting for Health in April 2007, with the on-line tutorial and testing available to staff free of charge.

United States of America

The American Medical Informatics Association (AMIA) and the national ICDL licensee, ICDL-US, worked closely during 2006 to create a US-version curriculum for 'anyone in a health-related entity who touches a keyboard containing person-specific health information' as well as an examination to certify mastery of this content. This will be entitled the Digital Patient Record Certification or DPRC. The curriculum group adapted the syllabus developed by the ECDL Foundation group; the US version was then reviewed by the ECDL group.

The test will be piloted in early 2007 with the expectation that the program will be functional late in 2007. ICDL-US and AMIA found the partnership to be mutually beneficial and there is a desire to work together on other products for the North American region.

AMIA has two major educational initiatives underway and this initiative is part of its "Got EHR?" campaign. AMIA also seeks to educate the general public about electronic health records and especially integrated personal health records, particularly as an integral part of the electronic medical record. The campaign also strives to increase the use of EHRs in small practice environments. The second initiative is the "10x10" Program, which is an effort to educate 10,000 applied clinical informaticians by 2010. This program now involves two universities and will involve at least five by the end of 2007.

Italy

The Italian Association for Computing and Automated Calculation (AICA), the Italian national ECDL licensee [11], instructed CERGAS Bocconi to work out the Italian health syllabus, based on the core syllabus, and the test structure. The related "ECDL Health Manual" has also been developed and printed. Between January and March 2007, Italy has implemented two pilot editions of the course, delivered to about 60 medical doctors and nurses of the Local Healthcare Units of Milan and Dolo (Venice). Participants have been offered four courses of ECDL Start (24 teaching hours), plus a specific course of ECDL Health (8 teaching hours). At the end of the pilot courses, in May 2007, final examinations will be held and skill cards issued accordingly (including, for the first time ever in Italy, those relating to ECDL Health). The examination to obtain the ECDL Health skill card will include practical exercises simulating the use of patient records management software. The courses will be included in national and regional programmes aimed at the continuous education of the NHS medical personnel and will enable participating medical doctors to obtain some compulsory education credits. These courses are expected to become a key element of the education and training programmes nationwide.

The Italian experience stands out for its special focus on the preliminary planning of the initiative, also being based on a scientific research project conducted in 2004 aimed at measuring the potential benefits of information education and training in the healthcare sector [12]. This research project has analysed and evaluated the "cost of IT ignorance" in the Italian healthcare sector through a sample survey, empirical measurement tests, and experience of a similar research project conducted on private businesses. Ignorance in the information field has proved to be a notable hidden cost for the Italian healthcare sector and the potential value of information education and training of the NHS personnel amounts to about 2 billion Euros per year. After being published and officially presented, the results enhanced the institutional awareness regarding development of targeted educational programmes.

Discussion

Development of any new qualification takes a considerable period of time, commencing with identification of the need. This initiative has sought to achieve this in a way that matches the differing needs of countries globally, and to pilot and validate it in differing countries in two continents. Though the piloting, and the development of the educational and testing frameworks and systems, was grounded on the needs of the individual countries involved, the resultant tools and products are likely to yield wider benefits and use by other countries wishing to follow suit.

As well as being a key contribution towards sustainable health informatics systems, this initiative also marks a policy development for the ECDL/ICDL Foundation. It began the development of the Health module as a global

concept with the help of a group of subject matter experts. However, it became clear that the application of Health informatics systems required a certification that was closely tailored to national requirements including patterns of practice, culture, language, and legislative frameworks. As a result the ECDL Foundation adopted a modem of product endorsement. The core syllabus for the Health supplement is specified by the Foundation, but individual national license holders propose the format in which the national certification is developed and assessed within a specific country, ensuring this has local relevance. Each national certification is then endorsed by the ECDL Foundation. The Health certification is thus tailored to each individual country's health system. It not only uses the languages of the country, but the terminologies and taxonomies for health care practice, as well as professional and legal codes. The ECDL/ICDL Health Supplement is the first successful implementation of this Endorsed Product concept.

Conclusion

With the increasing importance of health informatics systems, and the need to ensure their effective and safe usage, there has been a steady and increasing recognition of the importance of end user competence as a contribution to effective sustainable implementation and development of such systems. Moreover, given that the issues and risks are basically generic, coupled with increasing mobility of health staff, the advantage of devising a generic solution has become self evident. From this position, the progress in the last two years through an expert committee identifying and confirming the core syllabus, and three different countries undertaking trials and detailed implementation plans, is significant and encouraging.

Like motorcars, aeroplanes, or other items of advanced technology, health informatics systems are only as good as the competence (and confidence) of their users. Hitherto this has gone largely unnoticed, except for possible training in a particular system's operational instructions. The ECDL/ICDL Health Supplement has broken new ground, by recognising the high importance of the education and empowerment of the end user, whatever their level or profession. It thus makes a vital contribution to the sustainability of health informatics systems.

Moreover, this is a global solution, linking common generic requirements with local need through the ECDL Endorsed Product concept. Having been developed by an international expert group, endorsed by the ECDL/ICDL Foundation, and now validated simultaneously in three very different countries, this product is now available for use in any nation.

References

[1] Anderson JG, Aydin CE, Jay SJ (eds.). Evaluating Health Care Information Systems: Methods and Applications; Sage, Thousand Oaks CA, 1994.

[2] Rigby M. Improving Governance and Reducing Risk in Electronic Patient Record Systems: Ensuring Appropriate Competencies for Support and End-User Staff. In M Fieschi, E Coiera, Y-C J Li (editors). Medinfo 2004 - Proceedings of the 11th. World Congress on Medical Informatics, IOS Press, Amsterdam, ISBN 1 58603 444 8, 2004, 926-930.

[3] Rigby M. Protecting the Patient by Promoting End-user Competence in Health Informatics Systems – Moves towards a Generic Health Computer User "Driving License"; International Journal of Medical Informatics, 73, 151-156, 2004.

[4] Staff of EDUCTRA Concerted Action. Education and Training in Health Informatics: Guidelines for European Curricula; in Hasman A *et al.* Education and Training in Health Informatics in Europe; IOS Press, Amsterdam, 1995.

[5] International Medical Informatics Association. Recommendations of the International Medical Informatics Association (IMIA) on Education on Health and Medical Informatics; Methods of Information in Medicine, 39, 267-277, 2000. (also available on www.imia.org/wg1/rec.pdf)

[6] www.ecdl.com/main/index.php

[7] Rigby M, Draper R, Sheridan A. Confidentiality, Competence, and Confidence - Ensuring Ethics whilst Recognising Reality; in Proceedings, EuroRec '99 Third European Conference on Electronic Health Care Records, 1999, Sevilla, España; Sadiel, Sevilla, 1999.

[8] Rigby M. Protecting the Patient by Ensuring End-User Competence in Health Informatics Systems – Moves Towards a Generic Health Computer User "Driving Licence"; in Proceedings of IMIA Working Group on Education Conference, Portland, Oregon, 2003.

[9] Rigby M. The Health IT roll out: supporting the end users; Health Information Management, 33, 3, 93-94; 2004.

[10] Rigby M, Hulm C. Piloting in England of a Health Supplement for the European Computer Driving License (paper presented to HC 2007 and in press).

[11] www.aicanet.it

[12] Buccoliero L, Caccia C. Un approccio aziendale ai sistemi informativi in sanità. [A corporate approach to information systems in the healthcare sector.] In Borgonovi E, Camussone PF, Occhini G, (Eds.) L'Ignoranza informatica: il costo nella sanità. [IT ignorance: how much it weighs on healthcare] Milan, McGraw-Hill., 2004: 31-69.

Address of correspondence

Professor Michael Rigby, Professor of Health Information Strategy,
Centre for Health Planning and Management, Darwin Building,
Keele University, Newcastle-under-Lyme,
Staffordshire, ST5 5BG, UK.
Telephone +44 1782 583191, Fax +44 1782 711737,
email: m.j.rigby@hmp.keele.ac.uk

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

A Multi-method Approach to Assessing Health Information Systems End Users' Training Needs

Yiyu Qiu[a], Ping Yu[b] Dr. Peter Hyland[c]

[a,b] *Health Informatics Research Centre, University of Wollongong, Australia*
[c] *School of Information and Technology and Computer Science, University of Wollongong, Australia*

Abstract

Many existing training programs for Health Information Systems (HIS) do not adequately satisfy end users' needs nor meet training objectives. This is because they do not envisage the problems that users may encounter when performing specific tasks. Therefore the first priority for the development of an effective training program is to precisely assess the end users' training needs, a process called Training Needs Assessment (TNA). Applying traditional approaches for TNA, such as interviews or surveys alone, however, may be insufficient because they are limited in their capacity to reveal the cognitive processes of end users. Usability testing, with its ability to gather data about human computer interaction, overcomes the deficiencies of these traditional approaches. This paper proposes a multi-method approach, which combines usability testing together with traditional methods, such as interviews or questionnaire surveys to assess HIS end users' training needs. This innovative method is able to precisely reveal the training needs for different levels of HIS users. A case study, which applied this method to assess the training needs for users of a nursing information system demonstrates its feasibility.

Keywords:

health information systems, end users, training needs assessment, usability testing

Introduction

A well designed Health Information System (HIS) can not only save staff time in entering and retrieving client data, but can also increase the accuracy and completeness of such data. The adoption of HISs, however, is not common in the current Australian healthcare sector [1, 2]. One of the reasons is that shifting from the traditional paper-based documentation to electronic documentation requires the users of an HIS to have basic computer skills, to be familiar with the HIS, and to change their practices of information management [3]. Managers are more and more aware that end user training is an essential strategy to accomplish this [4, 5] and that the failure to implement appropriate end user training strategy will leave staff feeling frustrated and threatened by the new system [6]. In the worst case, they may even reject the new system or resign, which in turn will lead to loss of organizational resources like skilled healthcare workers. It is important, therefore, that significant investment be made into end user training and support in order to ensure that the introduced HIS will be accepted and used by the healthcare workers.

An effective end user training program should have the capacity to deliver a timely, effective, efficient and enjoyable learning experience to the end users [4, 7]. In other words, it plays the role of closing the gap between the complexity of an HIS and the users' cognitive capacity to master it. The majority of existing training programs, however, are not as effective as they promise [8, 9]. The most common problems of these training programs are that they are ill-directed and inadequately focused [10]. For example, some training programs provide healthcare professionals with huge amounts of unnecessary information because they have been developed as a "one size fits all" solution [8, 9]. The fundamental flaw is that training needs assessment, namely, the process of assessing the training objectives [11], is not properly conducted for these programs so the training program designer can not accurately envisage the problems end users may encounter when performing specific tasks using the HIS. Thorough TNA is required to improve end users' learning outcomes and to enable them to become familiar with a new HIS efficiently.

In order to conduct a thorough and accurate TNA, the method of analysis should be scientifically designed, and this is the topic to be discussed in this paper. Firstly this paper will critique the traditional methods for TNA. Afterwards it will propose a novel method for TNA, followed by a case study demonstrating how to integrate the new approach with traditional methods.

Need for new approach in training needs assessment

The primary purpose of a TNA is to identify what knowledge and skills end users should have in order to enable them to effectively interact with an HIS. Through identifying the usage problems, the gap between the necessary and the actual knowledge and skills that a user has for effectively interacting with this IS can be inferred [12]. In other words, experimental Human Computer Interaction (HCI) data such as usage problems, mistakes or inefficient behaviors, are effective indicators of what they do not

know [13], which, in turn, suggests what they need to know. Traditional methods of TNA, however, lack the ability to collect such detailed information about the cognitive process end users follow in their interaction with a new IS.

The common methods of assessing computer users' training needs are self-reported questionnaire surveys and interviews with end-users [13-15]. Questionnaire surveys have a number of distinct advantages, including the ease of distributing questionnaires to a large number of users and the automated analysis of the results with statistical packages. The typical process of quantitative assessment consists of managers setting the required level of skills for a particular task, then a staff member is requested to rate his/her skill level against this standard [16]. A comparison between these two sets of data suggests this staff member's skill gap in accomplishing the task, however, this method can only identify difficulties of which the designers or skilled external consultants are already aware [15]. It can not detect all the challenges or mistakes that a user may face or make in using a particular HIS.

Similar problems underlie other commonly used methods like interviews or focus group discussion, where end users are asked to reflect on their prior experience with an IS. Such qualitative assessments provide opportunities for users to express their perceived difficulties in using an IS but these conventional methods are not adequate to assess the learning needs of users, particularly for users with different levels of experience with the system. For example, novice users may not have sufficient knowledge about this new IS to enable them to realize the problems they may encounter in using it [17, 18]. Even experienced users may not be able to clearly recall their problems [19]. Verbalizing the process that a person follows to complete a task is also problematic, as it involves the expression of sequences of psychomotor movement in interaction with an HIS [20]. In addition, users' perceptions of the same problem may be different because of differences in their educational or technical levels [21].

Thus the use of traditional TNA methods alone may not reveal detailed information about the cognitive process end users follow in their interaction with a new IS [21, 22] so data gathered from these methods are not adequate to identify end users' training needs. In other words, basing the selection of training strategy on the analysis of such incomplete data may lead to ineffective programs because there is a significant gap between what is perceived to be a problem and the actual problems that a user encounters. Traditional methods for TNA need to be complemented by more effective new approaches. Careful observation of how users encounter and react to problems in interacting with a particular HIS is necessary. The next section will explain our proposed methods and suggest strategies for dealing with the problems discussed above.

Cognitive usability testing method

There is a growing role for the cognitive and behavioral sciences in health informatics, particularly as it pertains to human factors, and other areas such as information retrieval and educational research [22, 23]. From the perspective of informatics, cognitive science can provide a framework for the analysis and modeling of complex human performance in IS [23]. Theories and methods from the cognitive sciences can illuminate different facets of the design and implementation of IS [22, 23]. They can also play an important role in understanding and enhancing human performance on a wide range of tasks. These tasks may include developing training programs to reduce errors and increase efficiency for healthcare [23].

Usability testing encompasses a range of methods for identifying how users actually interact with a complete software product. Empirical testing is a form of usability inspection that observes actual user interactions with an interface (Neilson, 1994). Given that many HISs fail due to usability problems, organizations are starting to show interest in usability testing. Some preliminary studies have been performed testing clinical information systems. For example, Kushniruk [24] introduces a laboratory-based usability testing method to evaluate the effectiveness of HISs. According to Kushniruk [24], usability testing refers to "the evaluation of an IS that involves testing of participants who are *representative* of the target user population, as they perform *representative* tasks using an IS". During the testing, all interactions a user has with an IS are recorded (e.g. video recordings made of all computer screens or user activities and actions). In addition, this technique generally includes the collection of "think aloud" reports, involving the recording of users as they verbalize their thoughts while performing particular tasks [24].

In brief, this approach focuses on classifying users' cognitive ability and then identifying the problems they encounter during their interaction with the IS. With its ability to gather rich empirical HCI data, this method provides an excellent opportunity to complement the weaknesses of interviews or questionnaires for assessing training needs. Although this method was originally designed for testing the usability of an IS, it is also useful for identifying training needs. In the early '90s, Simpson proposed a framework to describe how testing methods could be used in the planning phase of designing online documents [25]. A recent case study has used novice users' interaction with a search engine to reveal the knowledge and skills that such users need [18]. Our preliminary study which employed the usability testing method to assess the training needs of nursing students also demonstrates the feasibility of this method [26].

A multi-method approach for training needs assessment

We have demonstrated the viability of cognitive usability testing methods in capturing process data on how an end user interacts with a particular IS. We propose that a multi-method approach combining usability testing with conventional methods like interview or questionnaires can precisely and thoroughly understand the process end users follow in processing information in a particular HIS. It can also suggest the knowledge and skills that these users need to learn in order to use the HIS.

The process of our proposed approach is to:

- observe how novice, intermediate and veteran users use an HIS to complete representative tasks.
- interview or survey these users before or after conducting cognitive usability testing, in order to ascertain their level of the knowledge and skills that are relevant to the HIS that they are trained to use
- extract patterns of strategies used by the novice, intermediate and veteran users to complete various tasks with this HIS.
- identify the key knowledge gaps (learning needs) of different levels of learners based on the different interaction patterns that they display when using the HIS.

In the following sections we describe three approaches which can be utilized for an innovative TNA.

Usability testing approach

The following three issues need to be considered when performing usability testing.

Participants

Participants for the experiment are potential or actual end users. Based on their experiences with the HIS (e.g. measured by the time and frequency of their usage of the system), they can be categorized as novice, intermediate, and experienced users. The recommended sample size varies between usability experts. As Nielsen and Mack suggest [27], usability testing can be carried out using only five participants, and the results will demonstrate 85% of the usability problems. Kushniruk *et al.* [21] suggest that up to 80% of usage problems can be detected from 8 to12 participants evaluating a system.

Outcome measurement

Kushniruk *et al.* [21] suggest that the usability testing should involve setting up recording equipment that allows for continuous recording of computer screens during the process of human-computer interaction. To achieve this goal, Camtasia Studio, screen recording software, can be used to record each participant's mouse movement and keyboard strokes. In addition, participants' "think aloud" reports can be audio-taped. The data analyzed for usability analysis included both the video and the audio file.

Data analysis technique

Prior to analyzing the video data, a coding scheme should be defined for use in identifying specific occurrences of users' problems and aspects of cognitive processes from transcripts of a participant's "think aloud" comments. A coding taxonomy developed by Kushniruk can be used for analyzing human computer interaction data [22, 24].

- Navigation: used when participants comment that they are navigating, or indicate that they are incapable of moving through the interface to find the relevant information or accomplish what they are supposed to do.
- Understanding: used when participants comment on understanding the meaning of labels, instructions or errors.
- Ease of use: used when participants comment on the level of "ease of use" of the system (from easy to hard) or any confusion or frustration experienced.

A case study

The NIS described in this case was the Care Planning Assessment Tool (CPAT), owned by the Hammond Care Group. The CPAT was introduced to help nurses carry out systematic assessments for nursing home residents. The "Clients" menu enables users to perform the most crucial functions of the CPAT, i.e., doing data entry and assessments for clients (i.e. residents), see Fig 1 for a screen shot of the assessment screen. In the assessment screen, users can enter detailed assessment results for a resident. The program can then generate various assessment related reports.

This particular study aimed to assess the training needs for different levels of the CPAT users, so as to develop the "right" training materials for the "right" groups of users. The training materials for the CPAT are:

- User Manual: this is usually for novice or first time users and should be very detailed.
- Online Help: this type of training is traditionally for the relatively experienced users who require help while using the product; it usually contains information on how to conduct a task.

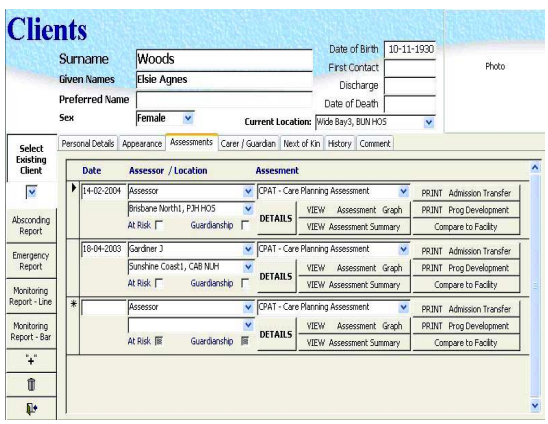


Fig 1 - The assessment screen of the CPAT

Procedure

Based on users' knowledge and experience with the CPAT, they are classified into two categories: novice and experienced users. The process of the training needs assessment in this case study was to:

- conduct laboratory-based usability testing to observe how novice CPAT users complete a series of data entry tasks using the CPAT;
- interview a cohort of experienced CPAT users to understand the problems they often encounter in using the CPAT software
- analyze two sets of data to identify knowledge gaps;

- integrate the findings into the design of training materials.

Usability testing experiment

Eight participants for the experiment were recruited from third-year nursing undergraduate students from the University of Wollongong. All the participants majored in geriatrics. They were potential users of the CPAT but had no previous usage experience with the software.

The participants were asked to perform the following three representative tasks supported by the software:

- entering data for a resident;
- doing an assessment for a resident;
- generating a change monitoring report.

The participants were encouraged to "think aloud" or verbalize their thoughts if they were uncertain about how to conduct the above documentation tasks using the software.

Findings

There were eight video and audio data collected from the experiment. Analyzing the triangulation of audio and video data identifies a series of problems that novice CPAT users encountered in this training session. An excerpt of a coded section of such triangulation is given below to show how users' interaction with the software was coded.

> 20:40 – user finished scoring "communication problems" and intended to answer the next group of assessment questions.
>
> "How should I go to the next group of questions?"
>
> *Navigation - having problems navigating between assessment criteria.*

By coding all of the participants' usage problems, three groups of problems that users encountered were identified:

- basic computer skills, e.g., users do not recognize the drop-down icon in the selection fields.
- knowledge about the software, e.g., users do not know where to score questions about assessment criteria
- domain knowledge about nursing documentation, e.g., users do not understand some assessment questions

Focus group discussion

A semi-structured focus group discussion involving nine veteran users from the Hammond Care Group was conducted to explore their learning and work experience with the CPAT. The participants were the actual users of the CPAT. Their roles included facility manager, trainer, consultant and dementia care worker. Most of them have more than three year's CPAT usage experience.

Participants were prompted to provide their answers for the following questions:

- How did you learn to use the CPAT?
- Could you recall any problems encountered when using the CPAT?
- What kind of help do you expect when you encounter problems?

Findings from the focus group discussion

The normal method of learning the CPAT was labor-intensive, one-on-one coaching, followed by self-directed practice. If the user encountered any problems, they could either approach the trainer or try to solve the problem by themselves. A list of frequent usage problems was identified through focus group discussion. It ranged from system-related problems to computer-related problems.

Discussion

The findings from the TNA identified both procedural and conceptual usage problems in the format of HCI (usability testing) and verbal expression (focus group), which provided valuable input into the design of both the user manual and online help for the CPAT software. For the user manual, usability testing vividly revealed three types of novice-user knowledge gaps, which would become the main focus of the user manual. In addition, frequently asked questions gathered from focus group discussion can be effective contents in the user manual as these problems are also common for novice users. For the online help, problems that were identified in usability testing and were summarized in focus group became the core contents of online documentations. In addition, feedback from experienced users suggested that the learners like functions like video demonstrations. Therefore, this new approach for TNA has proved its capacity to precisely and thoroughly identify the training objectives for both novice and experienced users.

Conclusion

In this paper we have firstly presented the strengths and weakness of a number of traditional methods (interviews and questionnaires) that are used for TNA for novice users of an HIS. Next we described the capacity of cognitive usability testing to capture the cognitive process of HIS users in their interaction with the system. We argue that using the conventional methods alone has limitations and that they could be complemented through combining cognitive usability testing with the conventional methods. This proposed new approach has been explained in detail, particularly through demonstration of a case study, which involved assessing the training needs for users of a nursing information system through two approaches: cognitive usability testing and focus group discussion.

Acknowledgement

We would like to acknowledge the contributions by:

All the participants in the laboratory and focus group.

Mr. Richard Fleming, owner of the CPAT application, for his support in providing introduction to the CPAT software for the participants of this experiment.

Dr. Patrick Crooks and Ms. Joanne Joyce and staff in Nursing school in University of Wollongong for encouraging and recruiting their students to participate in the experiment.

Dr. Madeleine Cincotta, for her generous help in editing and proof-reading the final manuscript.

References

[1] Gosbee J, Ritchie E. Human-Computer Interaction and Medical Software Development. Interactions. 1997;4(4): pg 13--8.

[2] Hunt EC, Sproat SB, Kitzmiller RR. The nursing informatics implementation guide. New York: Springer 2004.

[3] Yu P, Comensolia N. An exploration of the barriers to the adoption of information technology in Australia aged care industry. Health Informatics Conference 2004; 2004; Brisbane, Australia: HISA; 2004.

[4] Gjestland C, Van Slyke C, Collins R, Cheney P. End User Training: A Decade of Research; 1997.

[5] Huang AH. End-User Training Management Systems. Journal of Educational Technology Systems. 2002;31(1):19-33.

[6] Coiera E. Guide to health informatics: Arnold 2003.

[7] RadhaKanta M, Vincent SL. Evaluating end-user training programs. Association for Computing Machinery Communications of the ACM. 2005;48(1):n/a.

[8] Kidd MR, McPhee W. The 'lost generation': IT education for healthcare professionals. The Medical Journal of Australia. 1999;171(10):p510-1.

[9] Ward JPT, Gordon J, Field MJ, Lehmann HP. Communication and information technology in medical education. The Lancet. 2001;357(9258):p792.

[10] Washer P. Barriers to the use of web-based learning in nurse education. Nurse education today. 2001;21:455-60.

[11] Ostroff C, Ford JK. Assessing training needs: critical levels of analysis. In: Goldstein IL, ed. *Training and Development in Organizations*. San Francisco: Jossey-Bass 1989.

[12] Gupta S, Bostrom RP. End-user training methods: what we know, need to know. Proceedings of the 2006 ACM SIGMIS CPR conference on computer personnel research; 2006; Claremont, California, USA: ACM; 2006.

[13] Howard S. User interface design and HCI: identifying the training needs of practitioners. SIGCHI bulletin; 1995; 1995. p. 17-22.

[14] Nelson RR, Whitener EM, Philcox HH. The assessment of end-user training needs. Communications of the ACM. 1995;38(7):27-39.

[15] O'Brien E, Hall T. Training needs analysis: the first step in authoring e-learning content. 2004 ACM symposium on applied computing; 2004; Nicosia, Cyprus: ACM Press; 2004.

[16] Boydell T, Leary M. Identifying Training Needs. London: Institute of personnel and development 1998.

[17] Sein MK, Bostrom RP, Oflman L. Individual differences and conceptual models in training novice end-users. Human computer interaction. 1989;4(3):197-229.

[18] Lazonder AW. Exploring novice users' training needs in searching information on the WWW. Journal of computer assisted learning. 2000;16:326-35.

[19] Smith SM. Frustrated feelings of imminent recall. In: metcalfe J, Shimamura AP, eds. *Metacognition: knowing about knowing* 1994.

[20] Norton B. Selecting learning experiences to achieve curriculum outcomes. In: Billings DM, Halstead JA, eds. *Teaching in nursing : a guide for faculty*. St. Louis, Miss: Elsevier Saunders 2005.

[21] Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. Journal of Biomedical Informatics. 2004;37:56-76.

[22] Kushniruk AW, Patel VL. Cognitive approaches to the evaluation of healthcare information systems. In: Anderson JG, Aydin CE, eds. *Evaluating the organizational impact of healthcare information system: second edition*: Springer 2005.

[23] Patel VL, Kaufman DR. Cognitive Science and Biomedical Informatics. In: Cimino EHSJJ, ed. *Biomedical Informatics: Computer Applications in Health Care and Biomedicine (Third Edition)*. New York: Springer-Verlag 2006.

[24] Kushniruk AW. Evaluation in the design of health information systems: application of approaches emerging from usability engineering. Computers in biology and medicine. 2002;32:141-9.

[25] Simpson M. How usability testing can aid the development of online documentation. Proceedings of the 8th annual international conference on Systems documentation; 1990; Little Rock, Arkansas, United States: ACM; 1990.

[26] Qiu Y, Yu P. Nursing Informaion Systems: applying usability testing to assess the training needs for nursing students. Methods of Information in Medicine. 2006;to be appear.

[27] Nielsen J, Mack RL, eds. Usability inspection methods New York Wiley 1994.

[28] Courage C, Baxter K. Understanding your users: a practical guide to user requirements methods, tools and techniques 2005.

[29] Crowl TK. Fundamentals of educational research. Madison, Wisconsin: Brown & Benchmark 1996.

[30] Wilkinson D, Birmingham P. Using research instruments : a guide for researchers. New York ; London: RoutledgeFalme 2003.

[31] Morgan DL, ed. Successful focus groups: advancing the state of the art: Sage Publications 1993.

[32] Kirakowski J. Questionnaires in usability engineering. 2003.

[33] Kushniruk AW. Analysis of complex decision making processes in health care: cognitive approaches to health informatics. J Biomed Inform. 2001;34(5):365-76.

Address for correspondence

Yiyu Qiu
Email: yq518@uow.edu.au
69/39 Northfields Avenue, Wollongong, Australia 2500

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Making Health Informatics Competencies Useful: An Applied Health Informatics Competency Self-Assessment System

H.D. Covvey, S. Fenton, D. Mulholland, K. Young, B.A.E.

Waterloo Institute for Health Informatics Research, University of Waterloo, Waterloo, Ontario, Canada

Abstract

Years ago we undertook to define Health Informatics (HI) competencies. This effort resulted in the creation of a document that articulated HI roles, the challenges faced by HI professionals, the high-level tasks that they needed to undertake to address these challenges and the competencies (skills, knowledge, and experience) they needed to complete these tasks. Unfortunately, in so doing we created what is arguably the most boring book in history, shoes contents are very difficult to extract, use, maintain and improve. We report here the completion of a pilot of a system that we believe corrects this situation. It is a web-based tool that incorporates all of the material, from roles to detailed competencies, enabling them to be accessed and used for a variety of purposes, the most notable of which is professional self-assessment.

Keywords:

health informatics education, competency assessment

Introduction

Canada Health Infoway and other Canadian organizations involved in the deployment of ehealth infrastructure have pointed out that they face major challenges in accessing human resources with Health Informatics backgrounds. There have been estimates that as many as 2,000 positions in Canada go unfilled or are filled by less-than-fully qualified individuals.

Canada already has university and community college-based programs in HI, but these produce less than a hundred HI professionals per year. Even with a number of new education programs emerging, there is little hope of increasing this number significantly over the next 3-5 years.

In addition to this, the reality is that few students in our high schools or many of our post-secondary institutions have even heard of opportunities for careers in HI, or even what HI is!

This lack of awareness was echoed by academic experts and industry leaders from Canada, United States and Britain at the eHealth Conference 2005 in Toronto, Ontario who called for aggressive action to resolve the health informatics human resources gap. Rapid education deployment programs like AMIA's 10 x 10 program and the Waterloo Institute for Health Informatics Research Boot-camp program have been launched as attempts to correct this situation. However, there is still minimal awareness of the nature of HI, the roles that HI professionals perform, the education required to play a role in this field, and the opportunities for employment for graduates.

Moreover, many health-related public and private sector organizations that would benefit from HI expertise are only just becoming aware of the competencies required of health informaticians, or that education and training is available in our universities and colleges. One consequence is that recruitments are not properly informed regarding required competencies, when candidates are found, poor choices can easily be made in hiring.

The field of Applied Health Informatics (AHI), that most relevant to health organizations and the vendors that serve them, blends the informatics and health disciplines to find and deploy the best possible technical solutions to a wide range of healthcare information-related challenges. Thus, people in AHI need to have a solid technical background, a detailed knowledge of the healthcare industry and of the capabilities it has produced, and a wide range of personal, organizational and cultural, business, managerial and analytical skills.

AHI self-assessment system

We undertook several years ago to address the need for Health Informatics career awareness in part by beginning the development of a unique and innovative web-based Health Informatics competency self-assessment tool. This tool helps students, potential students, and the individuals who want to hire them: (1) to understand the types of roles that exist for individuals trained in Applied Health Informatics, (2) to understand the competencies required to fill these roles, (3) to self-assess their (or their candidate's) knowledge and skills against the competencies required in these AHI roles, (4) to take objective tests to validate their perceptions of their competencies and (5) to find educational resources that enable them to address deficiencies that are identified.

The primary purpose of this tool is to serve as a self-assessment system and to provide an index to learning resources. It is intended as a resource for students, teachers and employers related to understanding the skills, knowledge and experience expected of competent Applied Health

Informaticians. Furthermore, the competencies and competency categories embodied in the tool and supporting documentation[1] can assist curriculum developers in defining the educational content for Applied Health Informatics education programs. The system is also intended to be able to serve as a "front-end" or entry point to an educational program that can be used to document progress related to the competencies the student needs and has acquired.

It should be noted that our framework to define competencies and the tool we built to support self-assessment and access to learning resources are entirely general and can be applied to any discipline, not just to Health Informatics. This makes our framework and tool reusable and of broad potential interest.

Methodology

In earlier work [1], we led a process that defined HI competencies and supporting curricula using a team of approximately 100 health and Health Informatics professionals (including HI teachers/researchers, curriculum developers, human resources professionals in healthcare organizations and HI companies, vendor staff, government representatives, and current and candidate students). This work involved the comprehensive documentation of HI roles, the challenges faced by professionals in each of these roles, the high-level tasks (which we called "micro-roles") that professionals in these roles need to undertake to address the challenges, and the competencies required to accomplish the tasks/micro-roles [1,2]. This material was then used to define content that could be used as the basis for a number of programs world-wide [3]. One such program has been established at Conestoga College, in Kitchener, Waterloo, Ontario, which registered its first students in the fall of 2005. The competency definition project was funded by the Canadian Institutes for Health Research (CIHR), its products are frequently cited, and today it stands as one of the few definitions of HI competence that have been derived using a logical framework (a work breakdown structure) rather than being solely based on the preferences and opinions of teachers with their own programs.

Once having completed the documentation of competencies (there are on the order of 400 for AHI alone) and the other components of this project, we recognized that the form they were in needed to change. In particular, they were incredibly boring to read, making improving them an un-motivating task. In 2001, F. Lau at the University of Victoria made the suggestion that we consider some way of encapsulating or packaging the competencies and other content in a directly usable form. We considered this suggestion and realized that significant value could be derived by packaging the competencies in a software tool that made them accessible, available and integrated for review, understandable, and actually usable by individuals in the field.

We called this tool "WebSAT" and built it using the Web-based Informatics Development Environment (WIDE) developed by the Computer Systems Group at the University of Waterloo, led by D. Cowan [4]. WIDE is intended as a rapid pilot system definition environment, based on a decade of Software Engineering research, composed of software services and are customized using declarative techniques. The goal of WIDE is to reduce technological barriers to system design and development.

WIDE is primarily based on open source software technology and consists of a number of services and supporting frameworks. Applications can include input forms or reports containing extensive multimedia materials such as imaginative use of maps or any 2-dimensional diagram, websites, databases, indexing and searching methods, agents, and push technologies. WIDE also contains a knowledge management system that supports documentation of technical information and best practices. The structures underlying the services are usually expressed in an XML-based declarative language that uses metadata and XSL. In the WIDE metadata context, "programming" has effectively been replaced with a declarative methodology thus making it possible to provide a wizard or forms-based approach to building Web-based systems. Internally WIDE uses a bootstrap approach; its extensions are implemented using its own metadata technology. WIDE can support a rapid development paradigm and new applications can be quickly built and demonstrated. The components of WIDE are described next.

A mapping services framework. Interactive maps are delivered from a map-server, which supports zoom-in or zoom-out functionality and positioning over areas of interest. When connected to a database or other directory the maps can be used to display and interact with the location of a geo-referenced object. Map searches can be defined by a circle, or general polygon. The mapping service framework does not use traditional GIS software.

A diagram and chart services framework. This framework manages and delivers specified interactive diagram and chart types upon request for presentation of data on the Web or in other formats. The diagram and chart services are also based on the Scalable Vector Graphics (SVG) open W3C standard and so provide similar functionality to the mapping service framework.

An XML-based metadata framework. The structure of databases, websites, agents, and applications including reports and input forms with maps and diagrams are described using XML. They are transformed into operating applications through the use of XSL "programs." Any application can describe and subsequently access databases or Web sites reachable anywhere over the Internet.

A report services framework. This framework supports the management of interactive report and input form types including maps, charts and diagrams and delivers them on request for presentation on the Web or in other formats. The user indirectly specifies the form type and the data that that is to be presented or requested; the framework chooses the report or input form type and populates it with the requested data.

A content management services framework. This framework supports the management of text and multimedia information in a database where it can be viewed, searched, maintained and then published for use on the Web or in other formats.

An access control service framework. Access to any content such as a database, website or other text and multimedia content can be provided with multi-level access controls to determine who can read or change data.

A Web and database searching service framework. This service framework contains an indexing agent and search engine that will index known websites and databases and support searching. The results from Web searches are categorized based on different search criteria such as the proximity of words in a phrase. The results of combined database and Web searches can be presented together. The results of the two searches can be compared to see if new results have appeared in the intervening time interval.

A push/notification service framework. The general push/notification service framework allows developers to create systems that allow users to specify conditions under which they wish to be notified or have information pushed at them.

An agent service framework. The agent service framework supports the description of agents that will act autonomously to perform utility tasks within an application. Agents are often defined to manage redundancy. For example, agents could be defined to verify the content of "local" databases against authoritative sources or to allow a user to type information once while submitting the data to multiple databases or Web sites.

The academy – a knowledge management framework. The academy framework is used to support widespread dissemination of "documentation" and knowledge describing how applications can be built from the WIDE Toolkit.

Description of WebSAT

WebSAT (Web-based Self-Assessment Tool) is a Web-based AHI self-assessment system that enables individuals to review HI competencies, assess their own competencies and compare these to the competencies required for specific roles. The development of this system was funded using internal resources and volunteer labor and is now available as a demonstration website on request to the author. In return, we ask users to complete a brief assessment of the system for us.

Our competency definition work identified three types of HI professionals: (1) AHI (Applied Health Informatics) professionals, who define the requirements for, procure, deploy, implement, manage, guide the use of, and evaluate HI systems and methods in health enterprises and their supporting industries; (2) RDHI (Research and Development HI) professionals, who teach, do research, and develop innovative HI tools for the health system and are typically found in academia and private industry research labs; and (3) clinicians who need HI competencies to be good clinicians (called Clinician HI or CHI).

In order to make the initial version of the system of manageable scope, we limited ourselves to providing a tool for those interested in assessing their own competencies relative to those required by AHI professionals. Furthermore, we did not address many of the user interface issues. For example, we did not allow individuals to assess themselves first at a high-level (versus broad categories of competencies) and then at a more granular level within these categories.

This past year we extended the development of this tool through the able assistance of a student working in the Undergraduate Research Assistant (URA) program in the David Cheriton School of Computer Science (D. Chodos). This student added references to Web-based educational materials for sample set of competencies. This allows users to click on a link and access educational documents, on-line educational programs, and on-site courses that provide a means to correct competence deficiencies. This has turned out to be an excellent capability that supports lifelong learning in this field. This latter work was reported at the recent eHealth 2005 Conference in Toronto [5].

During the last 6 months we have made major improvements to the system:

- A new user interface has been developed that makes use of the system more intuitive.
- Users can now assess themselves at a high-level (compared to competency categories) or at a detailed level, reducing the effort if the user does not need to go to a deeper level.
- Improved graphics show the comparison of the user's stated competencies to those required by the selected role.
- There is a new query capability that allows recruiters to retrieve the competencies required for various roles, based on the importance of the competencies.
- The system supports the self-assessment of experience.
- A capability to take objective tests is now included so users can validate there perceptions of their competencies.

Finally, the overall system now has a better look and feel to improve the user experience.

Mode of operation

WebSAT operates as follows:

- Users who access the system are asked to register under a user identity and password that they create.
- The user can review the definitions and detail of any of the components used to define competencies, including possible roles, challenges, skills, etc. This is the part of the system that supports the understanding of AHI itself and the review of details that the content developer may add, delete or improve.
- The user can then select one of the potential AHI roles, after reviewing descriptions of each role. The users stated competencies will be compared with this role (this can be changed at any time).

- The user can then access approximately 20 "competency categories" (groups of similar competencies), each containing multiple specific competencies, and can assess him/herself as to the level of knowledge or skill he/she has. Competencies are assessable as one of seven levels from "no knowledge" to "expert"). These levels include: UNA=Unacquainted; ACQ=Acquainted; PAM=Passing Familiarity; GAM=General Familiarity; FAM= Working Familiarity; CAP=Capable; EXP=Expert)
- If the user wishes, she/he can drill down to detailed competencies within the competency categories.
- The system allows corrections and the saving of input for later completion.
- Once the user has responded to all competencies at whatever level of detail desired, the system compares the user's competencies to the selected role and provides graphical feedback (a bar chart of responses versus requirements) as well as textual feedback. Different roles can be tried, to see if there is a better fit with these.
- For areas requiring further work, the users have the opportunity to receive system guidance to educational resources.
- Users can access objective tests for each competency.

The competency categories included in the system are:

1. Personal Competencies for AHI
2. General Computing Competencies
3. Health Computing for AHI Professionals
4. Key IT Usage for AHI Professionals
5. General Health System Competencies
6. General Business/Management
7. General IS Department Management
8. Team and Human Resources Management
9. Re-Engineering and Management of Change
10. Strategic and Operational Planning
11. Assessment of the Value, Effects, and Cost of IT
12. General Technology/Systems Life-Cycle Management Competencies
13. Procurement Competencies
14. Systems Implementation and Integration
15. Systems Maintenance and Support System Customization/Ad Hoc Development
16. Project Management Competencies
17. Education and Training Competencies
18. Vendor/Service Provider Competencies
19. User and Process Observation and Assessment Competencies
20. Security Management Competencies
21. Information and Data Collection, Analysis and Management Competencies

A total of approximately 400 separate competencies allocated under these categories are addressed by the system.

Results

WebSAT has now been tested by both by graduate students in our HI program and by students in a new HI program at a nearby community college.

Students have reported a high degree of satisfaction with the system. They found the system useful and informative, and it gave them a clear view of what they still need to learn and where they stand relative to the requirements of various roles.

We have recognized, however, that other improvements are possible, and we continue to enhance the system further along the following lines:

1. We are undertaking further improvements to its user interface so that it can be more efficiently used by inexperienced users.
2. We are completing the addition of learning links to Web-based educational sources to be referenced by students with identified weaknesses.
3. We are adding additional objective tests and creating a new testing engine. The existing one does not support sufficiently complex multiple choice options.
4. We will extend the tool to incorporate Research and Development HI (academic-level) self-assessment to the same degree as the AHI assessment.
5. We are in the process of offering the system to other programs for use as a tool for students, and we are seeking recruiters to test its support for their activities.

Summary and Conclusions

We have developed a Web-based competency self- assessment system for Applied Health Informatics. This system allows individuals to assess themselves as to the congruence of their competencies with those required for specific roles. The system has been tested on students and is in the process of being enhanced and disseminated. Individuals interested in accessing the system personally or using it as a component of their programs are invited to contact the first author. The system is offered without charge.

Acknowledgements

This work was supported by NSERC under the Executive Industrial Research Chair program and by Canarie Inc. The authors are grateful for the contributions of D. Chodos, research assistant, at the University of Waterloo. The authors are grateful to F. Lau at the University of Victoria who first suggested the development of this system.

References

[1] Covvey HD, Zitner D, Bernstein RM. "Pointing the Way: Competencies and Curricula in Health Informatics", H/IT Advisors, March 2001 (Book: at http://hi.uwaterloo.ca, May 2007).

[2] Covvey HD, Zitner D, Bernstein R. "The Development of Model Curricula for Applied Health Informatics", Medinfo 2001, London, England, September 2-5, 2001.

[3] Covvey HD, Zitner, D Bernstein RM, MacNeill JE. "The Development of Model Curricula for Health Informatics",

Proceedings of the International Medical Informatics Association WG 4 Meeting, April 2003.

[4] Cowan D, Fenton S, Mulholland D. Lowering Technological Barriers to Developing, Deploying and Maintaining Web-based Interactive Information Systems, to appear in the International Journal of Technology, Knowledge and Society 2005.

[5] Chodos D. Accessing Applied Health Informatics Educational Resources via a Web-based Self-Assessment Tool, eHealth 2005 (Poster), May 2005.

Address for correspondence

WIHIR, DC3333, University of Waterloo, 200 University Ave. West, Waterloo, ON, Canada, N2L 3G1; dcovvey@uwaterloo.ca

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

E-learning for Students in their First Year: A French Experimentation at the Medical School of Grenoble

Jean-Marie Renard[a, d], Daniel Pagonis[b,c, d], Jean-Philippe Vuillez[c,d], Jean-Paul Romanet[c, d], Bernard Sele[c, d]

[a] *CERIM, CNRS-EA2694, Faculty of Medicine, Lille2 University, France*
[b] *SIIM, University Hospital, Grenoble, France*
[c] *TIMC, Faculty of Medicine, Grenoble University, France*
[d] *UMVF, French Virtual Medical University, http://www.umvf.org*

Abstract

A local study carried out in the Medical School of Grenoble shows that teaching in the first year in medicine studies satisfies neither the students, nor the teachers. The Faculty of Medicine of Grenoble decided to set up a reform in order to offer a high quality education. This reform leads to a complete reorganization of the curriculum and to the intensive use of new information and communication technologies of information, in particular, the use of multimedia documents. The communication and information technologies team of the Faculty of Medicine of Grenoble carried out an innovating and daring reform to start at the academic year 2006-2007. The new course is built on three activities: self learning on multi-media resources, meetings with teachers for questions-answers sessions and tutorials animated by older students. This article reports the first results for this successful project. In the academic year 2006-2007, are concerned 1290 students, 40 teachers and 8 disciplines.

Keywords:

e-learning; first year medical curriculum; multi-media; Grenoble, France.

Introduction

Observed difficulties during the first-year curriculum

A degradation of working conditions

This degradation was explained as a consequence of years of teaching without innovations or improvements. From a teaching point of view, the acquisition and transfer of knowledge remain extremely antiquated. The majority of learners remain passive. The student remains isolated before a task that cannot be managed without help. This may generate harmful consequences on his learning capacity..

A regular increase in the number of students

The increasing number of students is a factor which amplifies the degradation of the working conditions. Since 1975 the number of admission in second year has dramatically increased. The "numerus clausus" of Grenoble corresponds today to 12% of the number of registered students (year 2006, 1300 registered students against 166 students allowed to continue the second year course). The amphitheaters have a normal capacity of 650 places whilst 800 students actually attend the lectures. This number cannot possibly be increased for safety reasons.

A contest which no longer fulfils its role

The contest no longer fulfils its double role of regulating the number of students and selecting the best ones. Studies showed that the harder a contest is, the less effective the selection of the best students is. This strong selection has many direct and indirect consequences such as some influence on the content of the lessons. Difficult lessons only exist for their selective character and not for their relevance.

Inadmissible behaviours

The extremely selective aspect of the contest has generated an attitude of unfair competition behaviour, such as voluntary disturbance during the lectures by a few students, voluntary locking of the books to limit there availability, paper planes sent through the amphitheatres, going to the extreme: sale of already annotated lectures showing voluntarily added errors. These behaviours go against the university values of solidarity and equity amongst students.

Less equity between the students

Many a student registers in private courses which offer method and regularity in work in order to make up for the effects of these bad working conditions. This is also a means for the getting of annals and self assessment in relation to others. In the French system, the inscription to these courses generates a discrimination against money since the first access price exceeds 1500 € .

Effects over the forthcoming years

One of the indirect but considerable effects of this situation is that the motivation to work is getting slack. This is understandable considering the very great efforts provided for during the first year. Teachers estimate that students need three years to come back to a normal working attitude.

Solutions already tested by other universities were not very conclusive

Faculties of Medicine, facing the same difficulties, and drawing the same conclusions adopted various solutions:

- One of them is the repetition of the courses, which obliges the teachers to reproduce at various gaps the same course in many amphitheatres. That demands a great involvement from the teachers and availability of many extra amphitheatres (with a capacity exceeding 700 places). The risk is this might encourage the action of disturbing students, increase the loss of interaction between teacher and students and generate problems with regards to safety. For all of these reasons, this solution cannot be applied in all Faculties.
- Another very widespread solution adopted by various Faculties is the broadcasting of the course by Video-transmission in a secondary amphitheatre. The principle consists in diffusing on line and on large screen in a secondary amphitheatre the course done by the teacher lecturing in the first amphitheatre. This reduces the number of mobilized teachers. However, to be added further to the already listed problems, the high costs for expensive, fragile and not always reliable equipment which cause considerable discontent among students. Answering the student's questions is difficult and this increases the risk of disturbance.

The need for an important reform

Whatever the solutions considered by the universities, the quality of the teaching is unfortunately not improved, on the contrary it is damaged. To face the whole of these drawbacks, the Faculty of Grenoble team has proposed a reform which leads to an original solution.

Materials and methods

Elaboration of an original reform

Emergence of the project of a reform for the first year medical curriculum, all the actors at the medical school carried out a complete thinking on the modernization of teaching. Administrative people (Dean...), teachers, and students (elected representatives) endeavoured to find a solution which would improve the quality of the teaching. In there analyse, they took into consideration different dimension like behaviours [1,2], cognitive sciences [3,4] and possibilities offered by the communication and information technologies [5,6]. A reform combining a new organization of teaching associated with an intensive use of communication and information technologies appear to be the best alternative.

Specificities of the medical profession

Given the teacher's specific statute, the Faculty of Medicine offers more freedom. Most teachers of Medical School are not subject to a quota of hours before the students. A modification in manner of delivering of the courses will not be a problem due to their teacher statute.

Grenoble specificities

This project is in the continuity of previous reforms made for the years following the fourth year (2002-2003), and for the years between the second and the third year (2003-2004) of the medical curriculum. Another particularity is the strong motivation of all the teams involved in teaching for an intensive use of new communication and information technologies.

New methods of training

The reform does not change either the curriculum program or the organization of the contest.

The year (2006-2007) is divided into 2 six-month study periods. Each semester ends with a part of a contest (January and May 2007). One six-month period is divided into 12 training sequences, each sequence divided into 4 weeks.
Each week is devoted to a different training activity.

- The first week is devoted to the study of the courses on DVD-Rom: one or two different disciplines are studied over the week in the form of multi-media courses animated and wired for sound by the teacher's commentary.
- The second week is used for the formulation of all questions destined to teachers. These questions exclusively relate to the multi-media courses studied during the previous week. They are the base of the Question/Answer Meetings. The questions are formulated on the Faculty of Medicine of Grenoble's web site which is destined to the PCEM1 students (www.medatice-grenoble.fr). The on-line form is open each week, from Saturday to Tuesday exclusively. The access is secured by a login-password provided for at the time of final registration.
- The third week is devoted to the Question-Answer Meetings. Each discipline studied during the sequence is followed by a 2 hour Question/Answer Meeting. These meetings are held in small amphitheatres. Students are divided into eight groups. The meetings are ensured by the teachers in charge of the discipline. The answers are built from the questions collected on the on-line form (www.medatice-grenoble.fr).
- The fourth week is dedicated to the tutorials. The tutorials are directed by third year tutor medicine student. These meetings are intended for QCMs training. QCMs are validated by the teachers, corrected and explained during the meeting by the tutor students. These students are trained and supervised by the teachers in charge of the disciplines and the contest. The meetings are held the week following the questions-answer meeting.

Recording of the multi-media courses

Recordings are carried out by the Stendhal University Audio-Visual Service. The multimedia resources consist of animated slides commented and recorded by the teachers.

Evaluation of the reform

The evaluation of the project lies on statistical data describing the correct working of the project as well as on satisfaction surveys carried out weekly. At the end of each tutorial session, students are questioned. The information collected concerns the evaluation of the lesson and training methods.

Agenda for the project

The Table 1 reports the agenda for the project.

Table 1 - Agenda of the project

Date	Event
November 2005	Realization of the preliminary study
December 2005 to January 2006	Redaction of the performance specification
February 2006	Validation of the project by the Université Joseph Fourier's council
March 2006 to September 2006	Production of the multi-media records
July2006	Opening of the web designed for the students in their first year
August 2006	Production of the DVD-Rom number 1 (2000 pieces)
September 2006	start of the curriculum – distribution of the DVD-Rom N°1 to students

Results

Numbers of students and teachers

The teaching of the first year of the medical curriculum in Grenoble concerns 1290 students, 40 teachers and 8 disciplines for the year 2006-2007.

Self learning on DVD-ROM

There was 230 sessions corresponding to 460 hours of recordings. The slides are in flash-R, mp3 and xml format. The materials are equivalent to 220 hours of listening by the students.

Every six months, a new DVD-ROM is distributed to the students. The DVD-ROM contains 4, 7 Go of data (The mean size of one course is about 26Mo). DVD-ROMs are pressed in 2000 specimens. Each DVD-ROM cost 1,60 euro.

The courses are also available on the web site dedicated to students under the form of Podcast using Open source Xvid file format [7].

The user assistance for DVD-ROM usage only received 16 calls and all of them were solved by the assistance.

The hot-line was overloaded by calls coming from students registered in Universities other than Grenoble to get the DVD-ROM.

Students will fill an electronic form after each course. Every week 1050 students will fill the evaluation form. 36 000 forms will be delivered during the year.

The students' satisfaction rate ranges between 75% and 85%

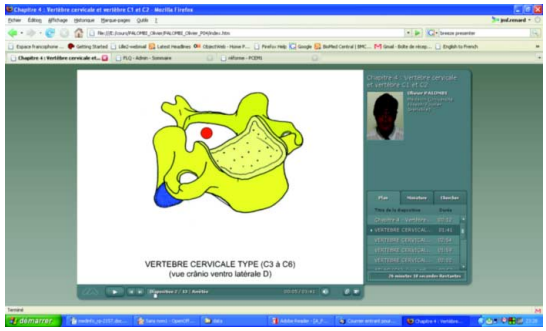


Figure 1 – self-learning rich media

The questions-answers sessions

The students send their questions through an on-line form. The questions must be related to the courses delivered during the previous week and are collected from eight independent groups of students.

At the end of the first cycle (the first four weeks), 900 students have used the system and asked 7484 questions which were marked "pertinent" 38 879 times by other students. The satisfaction level for that system was 85%.

The session with teachers was organized in small amphitheatres. Every student followed two sessions per week. Each session lasts two hours The ratio questions-answers sessions over self-learning time is two hours of questions-answers for ten hours of self-learning time.

The number of questions per teacher varies between 200 and 300. The quality of the question is qualified as good by teachers. The satisfaction level of the student is between 55% and 75%.

Tutorial session

The 1290 students are divided into 40 small groups. Every student follows two sessions per week. Each session lasts two hours. 96 hours per student and per year are envisaged.

The tutors are older students (120 students in their third year of medical curriculum). A session is animated by two students (always the same two students). The two students are supervised by teachers and spend two hours for the preparation of a session which content is a set of QCM.

1 000 000 QCM are planned to be written over the year.

The mean satisfaction rate is 94%.

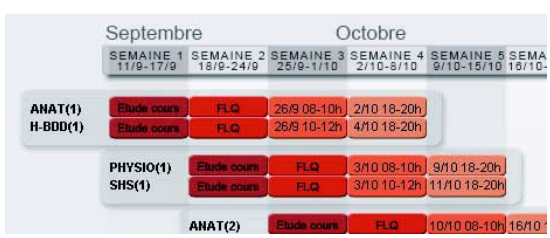


Figure 2 – The four activities of the learning cycle: self-learning on multimedia supports (Etude de cours), preparation of questions for teachers (FLQ), questions-answers session with teachers (ie 26/9 08h-10h) and tutorials (ie 2/10 18h-20h).

Logistics

A 700 places amphitheatre was converted into a multi-media room. It was adapted with 130 desktop computers, electrical connectors for personal laptop computers and with WIFI network connexion.

This room can accept 248 students who work in twos. It is open from 8h-20h and from Monday to Saturday

There is a low cost rental service for laptop computers and, in some cases, laptops can be lent.

Other rooms were adapted with computers and dedicated to questions-answers sessions as well as tutorials sessions.

During the first week, only 5 students used the 700 place amphitheatre. During the first month, the maximum number of students was 30.

The Web Site

It started in June 2006 and is reserved to first year students (www.medatice-grenoble.fr). It provides information on the organization of the curriculum. It gives access to the forms planned for their asking questions to the teachers. It is an entry point to access to the podcasts.

The average number of daily hits is 875 and the maximum 1200. In October, there has been 21 899 visits and 448 932 pages loaded.

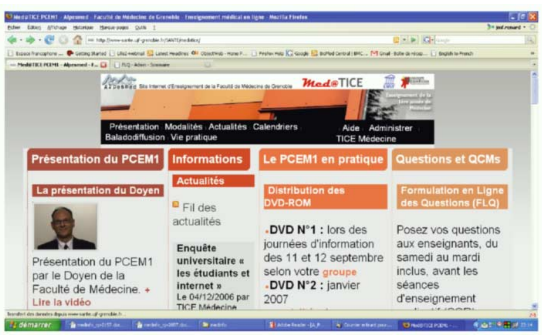


Figure 3 - The web site reserved to first year students

Discussion

The first result of satisfaction from students must be monitored to confirm the efficiency of this organization.

An element that needs to be confirmed is the uselessness of the multimedia amphitheatre.

A curriculum based on a four week cycle requires strong motivation from the teachers, how it will evolve in time is the question to ask.

The efficiency of a learning cycle lead on the sequence: self-learning on multi-média materials, questions-answers with teachers and tutorials with senior's student must be the subject for further analysis.

Conclusion

"The doctor of the 21th century must also have teaching competences in technologies which must be acquired at the time of his training" (Thierry KARSENTI).

The MedaTICE project showed that motivation and conviction of a team make it possible to carry out a joint and innovative project to it success.

Undoubtedly, the action taken by the Faculty of Medicine of Grenoble will be the starting point of a great scale disruption and will lead other faculties to become involved in similar projects.

It will be very important to follow on a monthly basis the statistics on the project and to do a complete analysis at the end of this first year.

Acknowledgements

Our warm thanks to Michèle Jacquemart for her assistance in preparing this paper.

References

[1] Apollon D. L'enseignement supérieur face au numérique en Europe : Tradition, innovation et résistance. Université de Bergen, 2006.

[2] Basque J. Influence de béhaviorisme, du cognitivisme et du constructivisme sur le design pédagogique.1999. www.teluq.uquebec.ca/~jbasque.montréal.

[3] Bernatchez PA. Vers une nouvelle typologie des activités d'encadrement et du rôle des tuteurs. Université de Montréal, 2001.

[4] Mayer, C. When learning is just a click away: Does simple interaction foster deeper understanding on multimedia messages?. Journal of educational psychology 2001: 390-397.

[5] Bernartchez PA, Marchand L. Encadrement médiatisé par ordinateur. Distances 2005 : Vol. 7. http://cqfd.teluq.uquebec.ca/distances/v7n1_b.html.

[6] Bellier S. Le E-learning. Editions Liaisons, Rueil-Malmaison, 2001.

[7] Xvid, Open source MPEG-4 Video Codec; Wikipedia Nov 2006. http://en.wikipedia.org/wiki/Xvid

Address for correspondence

Jean-Marie RENARD
CERIM, EA-CNRS-2694
1, place de Verdun
59045 Lille Cedex
mail: jean - marie . renard @ univ - lille2 . fr

Daniel Pagonis
Faculté de Médecine de Grenoble
Domaine de la Merci
BP 170 LA TRONCHE
38042 GRENOBLE CEDEX 9
mail: Daniel . Pagonis @ imag . fr

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

E-learning at Porto Faculty of Medicine. A Case Study for the Subject 'Introduction to Medicine'

Cristina Costa-Santos[ab], Ana Coutinho[c], Ricardo Cruz-Correia[ab], Ana Ferreira[bc], Altamiro Costa-Pereira[ab]

[a] *Department of Biostatistics and Medical Informatics, Faculty of Medicine, Univ. of Porto, Portugal*
[b] *Centre for Research in Health Technologies and Information Systems – CINTESIS, Faculty of Medicine, Univ. of Porto, Portugal*
[c] *Department of Informatics, Faculty of Medicine, Univ. of Porto, Portugal*

Abstract

The main objective of the Introduction to Medicine (IM) subject of the first year of the Medical Course at the Faculty of Medicine of the University of Porto is to provide students with a first contact with the areas of Biostatistics, Medical Informatics, Bioethics and the History of Medicine in the belief that they will be better prepared to learn, research, evaluate, share and decide within their practice. This paper presents a case study that describes how the subject IM is organized and how the b-learning tool (Moodle) is used to correct and grade the students' work. From the 239 students registered to attend the Introduction to Medicine subject 12% failed. The average grade among the successful students was 16 (out of 20). In the previous academic year only 2% of the students failed. However, among the successful students, the average grade was inferior (15 out of 20). The e-learning model that was described in this paper was successful because the results show that the students that made use of the Moodle got better grades.

Keywords:

education, medical faculty, e-learning

Introduction

The practice of Medicine requires the use of methods in order to acquire, store, process, analyse, transmit, evaluate and assess information as well as medical knowledge [1]. The study of scientific areas such as Biostatistics, Medical Informatics, Bioethics and the History of Medicine can contribute to improve the practice of Medicine and make it well supported by ethical principles and, therefore, socially effective. The main objective of the Introduction to Medicine subject of the first year of the Medical Course at the Faculty of Medicine of the University of Porto (FMUP) is to provide first year medical students with a first contact with the areas described above in the belief that, as future doctors, the students will be better prepared to learn, research, evaluate, share and decide within their practice [2]. Although there was some existing material on the Internet, it was only in October 1999 that an intranet for this subject was designed and implemented by the responsible professor and lecturers of the same subject [3]. The first interface, developed using PHP and HTML programming languages and a relational database, namely Postgres, in a LINUX server, was used to store and manage information. In the academic year of 2005/2006 an e-learning tool (i.e. Moodle) was introduced. This was done in order to use information systems' technology to provide for automatic students' work correction and evaluation.

This paper presents a case study that describes how the subject of Introduction to Medicine is organized and how the 'b-learning' tool is used to correct and grade the students' work in the academic year of 2005/2006.

Motivation for the use of an e-learning tool

A vast majority of first year Medical students are not interested in learning basic scientific subjects. They probably think that their content is not important for the clinical work ahead of them, and for which they have motivation. The introduction of technologies such as e-learning tools that bring new interactive methods for communication and simulation can help the lecturers to motivate the students and facilitate their learning process. Several case studies support the use of Web technologies in order to teach undergraduate Medical students [4].

Initial expectations

Although some of the material from previous academic years already included interactive means to communicate and learn (e.g. forums, chats, placard) the main objective to introduce Moodle was to use the specific module that allows students to make exercises and tests online, which are corrected automatically. With this module, the lecturers wanted the students and themselves to have a better understanding of the evolution of the students' learning results along the year with the main goal to improve the students' final results at the end of the year. This module eliminates the waiting time that was needed for the lecturers to correct the exercises and tests made by the students therefore simplifying the process of correcting and grading the weekly correction of 230 tests.

Objectives

The main objective of this case study was to create an e-learning platform that allows the students and lecturers to

follow the progress of students' work along the year. This helps the students to understand how they need to change or not their learning process with the real-time feedback they get.

Other objectives include facilitating students' access to the lectures' material as well as other interactive material and online communication means between students and lecturers that help and motivate the students to study and learn.

During the academic year, the lecturers noticed that the objectives were being achieved because the students were using regularly the material described above. This could be seen by consulting the usage statistics module within Moodle.

Materials and methods

Structure and grading system

Introduction to Medicine is a subject from the first year of the Medical Course of the Faculty of Medicine of the University of Porto that integrates four modules regarding the following scientific areas: Biostatistics, Medical Informatics [5], Bioethics and the History of Medicine. In the academic year of 2005/2006, 239 students were registered in order to attend this subject.

The Bioethics and History of Medicine modules are only taught with theoretical lectures and graded with a written examination. Each module weights 2 out of 20 for the final subject grade.

The other two modules, Biostatistics and Medical Informatics are taught with theoretical lectures, an e-learning component and practical lectures. Each one of these modules weights 5 out of 20 in the final classification (2 for the written examination and 3 for the work students do during the year at the practical lectures).

The 6 values necessary to complete the maximum grade of 20 come from a group project that is compulsory for all the students to complete at the end of the academic year. The group projects are supervised by the lecturers that help the students to understand and organize their work. In addition, students can attend seminars that give important information for the development and implementation of the group project.

Figure 1 - Moodle e-learning platform that is used within the Introduction to Medicine subject

B-Learning - integrating e-Learning components within the teaching process

The Biostatistics and Medical Informatics modules are taught with theoretical lessons (where the main concepts are presented to the students) and practical lessons (where the lectures support the students in their continuous self-learning process).

It is within the practical lessons that the integration of the theoretical concepts with the e-learning component available at the Moodle platform is made (Figure 1).

The practical lessons have the duration of 2 hours and a half every week and the students have access to their own personal computer. The lecturers start by discussing with the students the concepts given at the theoretical lessons and then these concepts are practised within the e-learning platform. The lecturers assess students' acquired knowledge (all the 239) individually with online mini-tests (Figure 2) at the end of every lesson. The mini-tests have a maximum duration of 5 minutes and include 2 or 3 multiple or numeric questions. These questions are selected randomly from a pool of questions that is created by the lecturers. The pool about a specific issue is big enough so that when students start the test they have most probably different questions to answer in their test or at least setup in a different order from all the other tests, so that copies among the students is almost impossible. All the information that is available within the Moodle can be accessed at anytime from any computer that has Internet connection. However, the mini-test can only be accessed during the practical lessons, only in some selected PCs where the lecturers need to insert a key so that the students can enter the right interface. This feature avoids problems with counterfeit of mini-tests because their results are taken into account for the students' final grade. When the 5 minutes expire the mini-test is corrected automatically and the final grade is showed to the student. If the student does not submit his answers before the mini-test finishes, the answers given until that moment are automatically submitted and the grade given accordingly.

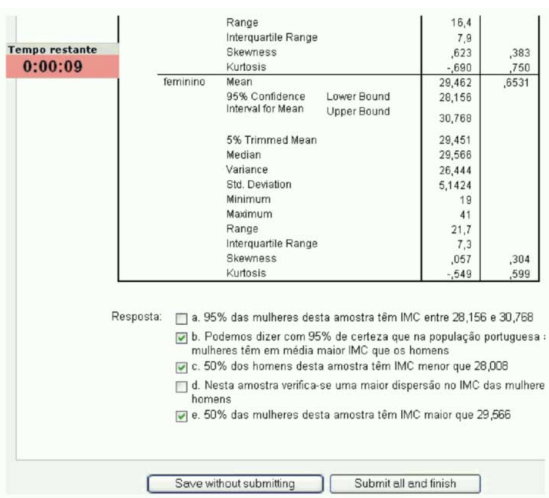


Figure 2 - Example of a weekly mini-test presented to the students, with automatic correction and timing

The average of all the results from the mini-tests can contribute to a maximum of 6 out of 20, which is the final grade.

Within the Moodle e-learning platform the content is organized on a weekly basis regarding each week of the academic year, in a total of 33 weeks.

Different types of content are included:

- SLIDES – the presentations made within the theoretical lessons;
- FORUM – used to exchange and share information about the subject between the students and the lecturers;
- EXERCISES – exercises to apply the theoretical concepts are made within the practical lessons and they allow the students to use real data from clinical studies. The tables with the real data that does not affect ethical issues related with the patients are also available in the Moodle.
- SELF-ASSESSMENT – the students can practice what they have learnt with self-assessment exercises with automatic correction;
- MINI-TESTS – these are exercises that count for the students' final grade. They are timed and corrected automatically by the Moodle;
- GROUP PROJECTS – the group projects are submitted by the students and commented by the lecturers within the Moodle, during 3 review phases;
- LINKS TO OTHER ONLINE MATERIAL:
  - MEDSTATWEB [6] - http://medstatweb.med.up.pt/ - the biostatistics interactive manual (figure 3) was developed by the lecturers and explains the basic biostatistics theoretical concepts with the use of practical examples and simulations (figure 4)
  - MEDICAL INFORMATICS - http://im.med.up.pt/ - the medical informatics manual was developed by the lecturers and explains theoretical concepts of Medical Informatics (see figure 5).

Advantages and disadvantages

Without an e-learning platform like Moodle it would be very difficult to make all the exercises and corrections that are done now because the number of students is too large compared with the number of lecturers.

This model allows students and lecturers to search and update the contents anytime, anywhere via Internet. It also permits the lecturers to monitor each student individually and the preliminary as well as final results of the subject and all the students in a generic way by allowing the analysis of the accesses that were made to its contents.

A disadvantage to consider can relate to the privilege that students with access to the internet at home may have in relation to the ones that do not have. To minimize this issue the students are encourage to use the computer laboratories within the Centre for Informatics at the FMUP that provide the Internet connection.

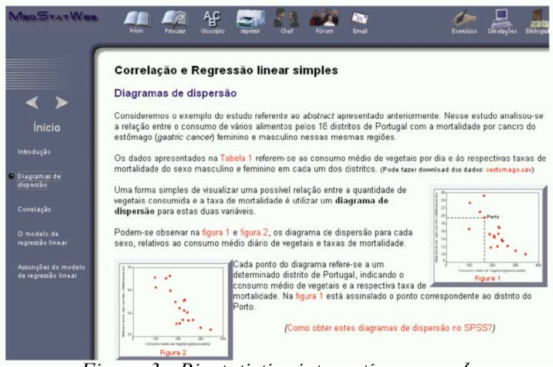


Figure 3 - Biostatistics interactive manual

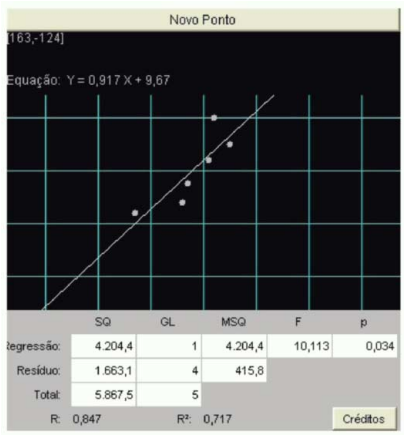


Figure 4 - Example of a simulation within the Biostatistics manual that explains the concept of regression. The user can move or insert points. The regression line is the ANOVA table and the relation coefficient automatically updated according to the points

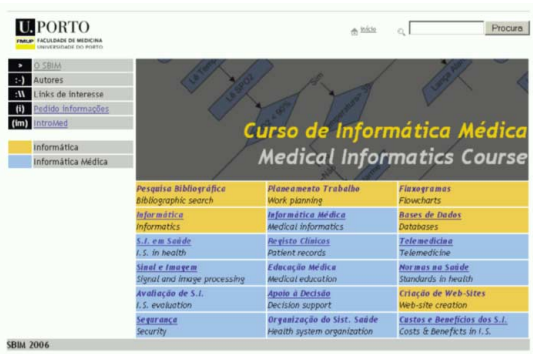


Figure 5 - Medical informatics manual

Structure and implementation

The implementation and integration of the Moodle e-learning tool was done in collaboration with the Centre for Informatics at the FMUP. The contents have been developed, improved and updated by the lecturers during the past 10 years. In the academic year of 2005/2006 a group including 7 lecturers, the professor responsible for the sub-

ject and 2 other people from the Centre for Informatics at the FMUP integrated the contents of the subject within the Moodle platform, version 1.5.3 [7] along with a web server and a Mysql database server version 4. Both servers are shared with other services that the FMUP provides so the e-learning system does not have a specific server adjacent to it. Nevertheless, the performance of the system is not compromised. Both tools (Moodle and Mysql) are open source software so they do not bring additional costs to the FMUP. Furthermore, Moodle is a flexible tool, and being open-source software, it can be altered in order to adapt some of its functionalities in order to generate data and lists of relevant statistics to be used in this case study. The actual version of the Moodle does not include these functionalities.

The Moodle platform is available at http://moodle.med.up.pt.

The authentication platform to Moodle is synchronized with the authentication for the users of the FMUP allowing for its easier integration.

This platform was also used by integrate another subject form the Medical degree and 5 more subjects from a post-graduate summer course organized and lectured by the Biostatistics and Medical Informatics Department.

Results

Results for the use of the biostatistics module

The biostatistics module was lectured with theoretical and practical lessons during the first semester of the academic year of 2005/2006. During the course of this semester 227 students accessed the e-learning material available from Moodle. The students made an average of 1,4 accesses to each chapter of the interactive manual, the MedStatWeb, 1,5 accesses to each exercise of the practical component, 2,8 downloads of each set of slides from the seminars and 2,5 downloads of each set of slides from the theoretical lessons.

From the 227 students, 31 got an average grade inferior to 10 (out of 20) in the mini-tests, 149 got an average grade between 10 and 15 whilst 47 students got grades equal or superior to 16.

202 students wrote the final exam and the average grades of the Biostatistics module was 11.

We noticed that the students with better grades in the mini-tests were the ones that made more accesses to the MedStatWeb interactive manual and practical exercises within the Moodle. Although some of the students with better grades also made more downloads of the theoretical lessons' slides and seminars, these differences were not statistically significant (table 1).

Evaluation of papers

All papers will be reviewed by three (3) independent reviewers. Reviewers will use a standardized form for review, a sample of which appears on this Web site (click here for direct link to that form). Authors should be familiar with the criteria used in evaluation.

The students that got an average grade inferior to 10 in the mini-tests for this module got an average of 7 (standard deviation of 4) in the final exam for this same module. Those who got an average grade between 10 and 15 in the mini-tests got an average of 10 (standard deviation of 4) in the final exam. The students that got an average grade equal or superior to 15 in the mini-tests got an average of 14 (standard deviation of 4) in the final exam. These differences have a statistic significance (p<0,001 - One-Way ANOVA).

Table 1 - Median (minimum-maximum) of accesses made by the students for each type of content available in Moodle

	Average grades for the weekly mini-tests			
	<10 n=31	**10–15** n=149	**>15** n=47	**p***
Number of accesses during the first semester to:				
7 chapters from the interactive manual	8 (0-36)	10 (0-38)	13 (0-51)	**0,047**
10 exercises with real data	15 (6-25)	15 (4-29)	16 (5-33)	**0,033**
10 sets of slides – seminars	24 (1-136)	25 (0-125)	30 (2-66)	0,283
10 sets of slides – theoretical lessons	23 (2-91)	24 (0-73)	27 (5-56)	0,386

* kruskal Wallis test

Results for the use of the medical informatics module

The medical informatics module was lectured with theoretical and practical lessons during the second semester of the academic year of 2005/2006. During this semester 221 students accessed the e-learning material available in Moodle. The students made an average of 2,2 accesses to each exercise of the practical components, 2,5 downloads to each set of the theoretical lessons. From the 221 students, 37 got an average grade inferior to 10 (out of 20) on the mini-tests, 124 got an average grade between 10 and 15 and 60 students got an average grade equal or superior to 16. 202 students did the final exam for this module and the average grade was 12. The students with better average grades in the mini-tests made significantly more accesses to the practical exercises than the ones who did not. Although some of the students with better grades also made more downloads of the theoretical lessons' slides and seminars, these differences were not statistically significant (table 2).

The students with an average grade inferior to 10 in the mini-tests within this module got an average grade of 10 (standard deviation of 6) in the final exam on the same module.

Those who got an average grade between 10 and 15 in the mini-tests got an average of 13 (standard deviation of 5) in the final exam. The students that got an average grade equal or superior to 15 in the mini-tests got an average of 15 (standard deviation de 3) in the final exam. These differences have a statistical significance (p<0,001 - kruskal Wallis test).

Statistical data for the use of the platform

During the whole academic year the students made, within the Moodle, about 22 000 accesses to the slides used for

Table 2 - Median (minimum-maximum) of accesses made by the students for each type of content available in Moodle.

	Average grades for the weekly mini-tests			
	<10 n=37	**10–15** n=124	**>15** n=60	**p***
Number of accesses during the second semester to:				
6 exercises with real data	11 (2-23)	13 (0-38)	14 (6-45)	**0,026**
9 sets of slides – theoretical lessons	17 (0-39)	22 (0-77)	25 (1-60)	0,070

* kruskal Wallis test

theoretical lessons and seminars, 6 650 accesses to the practical exercises that relate with the theoretical concepts, 3685 downloads to real data databases, 10 383 submissions of the self-assessment exercises with automatic correction, 3 947 mini-tests that were included in the student final grade, 420 group project submissions and 94, 29 and 14 discussions made within respectively the forum shared between students and lecturers, the forum used only by the lecturers and the forum where news were announced to the students.

The e-learning model developed by the lecturers supports and motivates the study using other web material such as interactive manuals namely MedStatWeb and Medical Informatics manual.

There was an average of 53 daily visits was made to the MedStatWeb, with 2565 hits per day. Each student made an average of 6 visits per month on this same manual. There was an average of 38 daily visits to the Medical Informatics interactive manual with a total of 587 per day. Each student made an average of 5 visits per month on this same manual.

Anonymous survey applied to the students

At the end of the academic year of 2005/2006 anonymous surveys were completed by the students that attended the subject of Introduction to Medicine and came to write the final exam (n=170). 90% of the students agreed that there was a good relationship between students and lecturers, 82% that the subject was well organized, 79% that it was easy to access the available material, 75% believe that was motivated by the interaction of the practical lessons, 65% thought that was able to pay attention and be interested during the lectures, 59% said that they accessed the online material on a regular basis, 54% stated that they could study regularly the subject and 36% said they attended the theoretical lessons regularly.

Final results

From the 239 students registered to attend the Introduction to Medicine subject, 231 were assessed from which 204 (88%) were successful and 12% failed. The average grade among the successful students was 16 (out of 20). In the previous academic year, when another e-learning platform was in use only 2% of the students failed. However, among the successful students, the average grade was inferior (15 out of 20).

Conclusions

The e-learning model that was described in this paper was successful because the results show that the students that made use of the Moodle got better grades. Also, the results from the survey illustrate that a vast majority of students think that the subject I well organized and that was easy to access the available study material. They also agree they were motivated to participate more actively in the learning process.

It is important to refer that the significant statistic differences reflect the fact that students that access more the online exercises and simulations are the ones that got better grades, not the ones that made most downloads of the lectures content material.

However, it is surprising that the students made so many downloads from the slides available since one download and printout for each set of slides would be enough.

According to the results obtained the mini-tests are beneficial for the students learning process. The students that could not obtain a satisfactory grade in the mini-tests could not obtain a similar grade in the final exam. On the other hand, the students that got good grades in the mini-tests were able to get the best results in the final exam.

This e-learning platform facilitates and simplifies the lecturers work because it would be very hard for them to correct 200 mini-tests every week and give the students the updated feedback every time. The integration between the traditional teaching with the e-learning was successful as 90% of the students agree that there was a good support between them and the lecturers and this happened mostly during the practical lessons.

In conclusion, although the percentage of failed students was bigger than in previous years (reflection of a tougher final exam because lecturers thought the students were better prepared this time), the final grades obtained for the Introduction to Medicine subject were very satisfactory because the students that passed the final exam had generally better grades than the ones in previous years.

For the future use of the Moodle platform the layout needs to be improved to be better adapted to the FMUP characteristics and objectives. Also, the use of forums and discussion groups will be more encouraged as they can help the students to clear their doubts and discuss the lecturers' topics in a more informal way.

Acknowledgements

We would like to thank the Introduction to Medicine lecturers for the academic year 2005/2006 that are not co-authors for this paper but participated actively in this e-learning project, namely Luís Filipe Azevedo, Filipa Almeida, João Fonseca and Mário Ribeiro. Further we would like to thank other collaborators such as: Camila Dias, Clara Tavares and Pedro Marques, as well as Armando Teixeira Pinto and Alberto Freitas that did not lecture this specific academic year but have contributed in previous years to improve the use of the e-learning tools. Finally, special thanks to Miguel Ângelo from the Centre for Informatics at the FMUP that contributed with the implementation of the Moodle layout.

References

[1] Costa-Santos C, Ayres-de-Campos D, Costa-Pereira A, Bernardes J. Utilização do Web Site de um Projecto de Investigação na Educação Médica. Educação Médica 1998; 9 (2): 110-6.

[2] Costa-Santos C, Caso de Estudo de Introdução à Medicina. Elearning UPorto 2004-2005 : 331-340.

[3] Costa-Santos C, Teixeira-Pinto A, Costa-Pereira A. Web-based Learning at the Faculty of Medicine of Oporto: Description and Evaluation of a Student's Internet. Proceedings of the 8th International Conference of European University Information Systems 2002; 152-4.

[4] Bernardo V, Ramos MP, Plapler H, De Figueiredo LF, Nader HB, Ancao MS, Von Dietrich CP, Sigulem D. Web-based learning in undergraduate medical education: development and assessment of an online course on experimental surgery. Int J Med Inform. 2004 Sep;73(9-10):731-42.

[5] Freitas JA, Cruz-Correia R, Almeida F, Costa-Pereira A. Evolution of Medical Informatics teaching in a medical undergraduate course .Technology and health care - official journal of the European Society for Engineering and Medicine: Proceedings of MedNet 2003, Eighth World Congress on the Internet in Medicine 2003; 11 (5): 312-314.

[6] Teixeira Pinto, Parry G, Vieira P, Martins C, Costa-Pereira A. MedStatWeb: a Web-Based Course on Medical Statistics. Proceedings of the meeting Web-Based Learning Environments 2000:158

[7] Moodle - a free, Open source Course Management System for Online Learning. Available at http://moodle.org/

Address for correspondence

Ricardo João Cruz Correia (rcorreia@med.up.pt)

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Designing m-Learning for Junior Registrars – Activation of a Theoretical Model of Clinical Knowledge

Anne Marie Kanstrup[a], Niels Boye[b], Christian Nøhr[a]

[a] *Virtual Centre of Health Informatics, Dept. of Development and Planning, Aalborg University, Denmark*
[b] *Regional Hospital of Randers / Aarhus University Hospital, Denmark*

Abstract

The MINI-project aims at supporting junior registrars in the learning process of how to utilize their theoretical knowledge from Medical School in everyday clinical reasoning and practice. Due to the nature of the work – concurrent moving, learning and producing - we designed an m-learning application. This paper introduces the possibilities and challenges for design of the m-learning application based on a) analytical findings on learning and mobility as derived from the design case – an emergency medical ward b) theoretical perspectives on medical knowledge, and c) presentation of the design of an m-learning application. The design process was based on user-driven innovation and the paper discusses considerations on how to combine user-drive and generic models.

Keywords:

software design; user-computer interface; medical informatics applications; decisions support systems, clinical knowledge; physician practice patterns; clinical reasoning; diagnostic reasoning; handheld computers; guidelines; education; internship

Introduction

Handheld computers or personal digital assistants (PDAs) have become common in clinical settings, and are used by physicians as well as nurses [2; 6]. The most common functions for PDAs in clinical settings are to provide clinical reference guides for drug information, patient tracking and various clinical guidelines. Increasing wireless connectivity combined with more patient data in digital form, introduces new application areas such as electronic prescribing, real-time medical records access and point-of-care evidence-based literature searches [13]. This paper reports from the MINI-project which is experimenting with the use of handheld computers for mobile e-learning/ m-learning to support physicians' introduction to clinical work. The case is an emergency medical ward at a large regional hospital in Denmark.

The case: learning and mobility at a medical ward

During their first years of medical practice physicians need to operationalize their knowledge from medical school, in the terms of Dreyfus & Dreyfus going from "knowing that" to "knowing how" in stages from novice to competent [5]. The current practice for supporting physicians in this learning process is an apprenticeship process [10] where junior registrars make experiences in daily hospital production and 'reflections on action' with chief physicians. The 'reflection *on* action' [14] is organized formally in conversations taking place during shifts, x-ray and other conferences but can alternatively, if the problem justifies it, take place "on the fly". At these conferences physicians (junior registrars and chief physicians) discuss their experiences, diagnoses, the situation and status of patients at the ward, care plans, etc. When it comes to 'reflection *in* action' [14], the back-up of registrars and conversations between novice and experienced physicians is, however, more difficult especially due to time pressure and lack of resources. Photograph 1 displays the ward used as case and is taken during ward-rounds displaying a typical situation of contact between junior registrars (standing in a circle) and an experienced chief physician, here briefly passing by answering questions and giving suggestions and then on to the next patient. At a design-workshop focusing on learning resources when on duty, junior registrars expressed how "you are mostly alone when on duty and when making your decisions" and how "one of the difficult things to learn is to make decisions on your own". Their back-up for decision-making is carried in the pockets of the white-coat, which is stuffed with reference books, instructions and guidelines, personal notebooks, etc. as seen in photograph 2.

Photograph 1 - Chief physician giving brief advice to registrars on duty. Photograph 2 - The pockets/back-up of a registrar (2,6 kg).

The objective of the MINI-project is to experiment with design of m-learning [11] to support physicians especially in their first period of clinical work as junior registrars. The aim is to design ePockets to replace the paper pockets displayed in photograph 2 and more importantly to improve knowledge support for decision making by inexperienced but professional physicians: the right information, at the right time, at the right place, and right at hand via mobile technologies and m-applications.

Empirical work at the ward/with staff (observations, interviews, conversations, workshops, design-meetings) has pointed out the following important characteristics for learning and mobility in this specific design context [7]:

- Mobile means walking – the physicians walk while looking up information and interaction techniques for 'one the move-interaction' like one hand interaction, is important.
- Mobile means on the way to the next task – maximum time for looking up information or making notes is 3 minutes and navigation supporting 'easy to find' is important.
- Learning at work is filled with interruptions – the physicians are constantly interrupted and as a consequence working hard to keep track via personal notebooks on their patient. It is important not to design yet another interrupting artifact but rather a personal device supporting their tracking of their work.
- Learning takes place in situ when working. Consequently it is important that the learning content is related to this 'situ' (vs. general information from reference guides, etc.)
- These characteristics of mobility and learning challenge the dominant 'office-domain' and call for experimentation with interaction techniques and content development (e.g. from text towards AV or VR).

Methods: user-driven innovation and (generic) models

The primary outset for the design process has been user-driven innovation in order to root the design in its context [8]. The sustainability of the specific design is to be seen in this perspective – an information ecological perspective to design and use of information technologies as an interwoven part of practice, values, people and technologies [12]. User-driven innovation is a meeting-point for designers and users making it possible to swap roles and bring use-practice, values and users to the foreground of design-processes. Hence, user-driven innovation provides a nice point of departure for sustainable software development from an information ecological point of view.

To combine design and use, theory and practice, is, however, not an easy task solved by involving users in the design process. A significant but very difficult aspect of software development in a wider perspective than just one case (one information ecology) is the development of generic models which make it possible and easy to make, share, maintain, and not least understand applications. Experiences with development of Electronic Health Records (EHR) emphasize this difficulty where the concept of archetypes has been elaborated for many years and where the goal has been sustainable systems with an easy exchangeable knowledge component [3]. Currently, a similar picture is surfacing within the area of providing online instructions or guidelines to medical staff in Denmark. Several different projects are emerging most of them focus on transmitting information to the WWW or PDAs. In the MINI-project we have developed a theoretical model for clinical knowledge as the important basis for understanding the problem to be supported and not least as a basic model for the designed m-learning application. The following presents this model and how it has been activated in the design of our MINI-m-learning application.

Towards a model of medical knowledge

Our employed model of knowledge in the medical domain is inspired from Adolfsen's epistemological model of everyday problem solving, where knowledge is stratified in three layers [1]. We define medical knowledge as consisting of three interrelated layers.

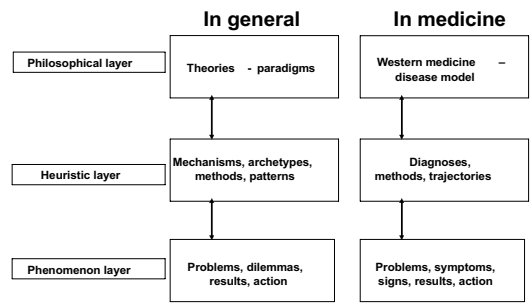


Figure 1 - The employed model of knowledge. The western medicine model is simple and says: "That to every disease or syndrome is a specific lesion, this can be functional, anatomical, physiological, biochemical, genetic, or social". Another model in the philosophical layer is Watson & Cricks model of DNA. For the other parts of the figure, see text for explanation.

The figure shows how the same dilemma can co-exist in three appearances of abstraction: in the bottom level the concrete everyday problem in contact with the everyday action and environment. In the middle heuristic layer, as a general problem, an archetype, a pattern, a diagnosis, or a method. In the upper model layer, as a theory or philosophy.

A medical example:

- bottom level – the patient's problem: palpitations corresponding to the doctors problem in this phenomenon layer: tachycardia (fast heart-rate)
- In the heuristic, archetypical layer the disease entities that can cause fast hearth-rate are located: e.g. thyrotoxicosis, fever or heart disease (there are many more).

- At the theoretical level are factors promoting a fast depolarization of the sinus node in the heart.

There are multiple other causes of fast hearth-rate than given in this example, this serves as an illustration of the different instantiations of the same problem and that "clinical reasoning" is translation and knowledge acquisition, activation, and operationalization (only possible in the phenomenon-level) in the different layers and this is not a strait forward algorithmic exercise. A recent review article "Educational strategies to promote clinical diagnostic reasoning" discusses this in detail mainly in the phenomenon layer [4].

Doctors need to master a problem in all three layers including the action relevant to diagnostic and therapeutic procedures. Patients have their primary interest and attention in the phenomenon layer, since the concrete problems are experienced in or on her body. Patient education efforts aim at bringing patient knowledge and reasoning to the heuristic layer, and thus provide them with "higher level tools and understanding" in cope and self-care of their disease problems.

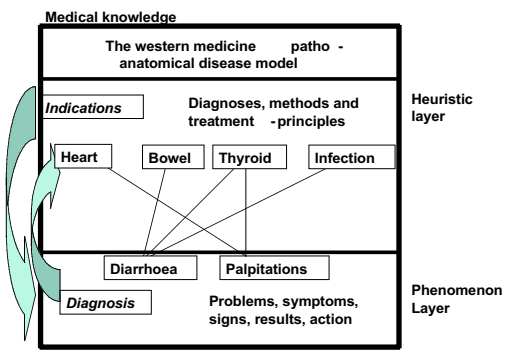


Figure 2: Example of the (non-algorithmic) dynamics of operationalization of medical knowledge. A patient presents diarrhoea and palpitations as problems. The doctor generalizes these problems to diagnoses in the diagnostic process. During this he activates knowledge from the upper theoretical layer. After the diagnostic process an indication of treatment is formulated employing both theoretical and heuristic knowledge about treatment principles. The principle is brought into action tailored to the specific problems and patient. The explanation for the choices made is "the indication" for the treatment.

Designing m-learning for medical knowledge

This perspective on medical knowledge put emphasis on support for the bridging between the different layers of medical knowledge – the relations – when designing m-learning for physicians in medical training i.e. focus on the heuristic layer as the meeting point for relations and support knowledge to and from this point. This perspective and focus is not new in the practice from the medical ward used as case. The heuristic layer is currently supported by the ward-instruction "the acute" organized in archetypes (heuristics for the ward e.g. 'the patient with fever, 'the arthritic patient', etc.) with relations to both the phenomenological layer (e.g. symptoms) and less prominent the philosophical layer (from University studies).

Consequently, based on both empirical analysis and the theoretical perspective on medical knowledge, the point of departure for the MINI-project has been to digitalize archetypes a) because it is learning content related to the specific use situation of the m-learning application b) because it supports the relations in medical knowledge grounded in the heuristic layer. It implies that the junior registrars are respected as professionals able to perform "clinical reasoning" and activate different layers of relevant knowledge regarding the specific patient – to fit a patient into a correct archetype and perform the necessary individual patient specific adjustments from this non-existing, average patient described in the archetype.

The transformation of existing paper archetypes into digital archetypes was carried out on the basis of a modified general work-flow model: with four "milestones" – all situated in the phenomenon layer: conclusion of the diagnostic interview (anamnesis, symptoms and signs) – often the basis for activating a specific archetype

1. Checklist of symptoms and signs contained specific ideas for further and alternative information relevant to a more precise positioning of the patient within the range of the archetype diagnoses.
2. Checklist of diagnostics, objective findings and test – contained relevant ideas for specific hospital procedures.
3. Checklist – treatment-plans
4. Checklist – monitoring and alternative actions.

In the technical dimension transformation of the paper archetype aimed at:

- A structure useful for a database
- Software useful as editor for producing archetypes within this structure
- A navigation design for MINI-archetypes/PDA-interaction

Microsoft Word™ was used as editor for production of the text for digital archetypes. Microsoft Word™ was used for producing the existing text for ward instructions and therefore well-known software to the archetype-producers (staff specialists). The archetype-producers were allowed to mark and prioritize the order in which they entered the text for later hyperlinking by means of a parser constructed by the programmer in the MINI-project.

The mini-mizing of the rather long text for each archetype (some were up to 25 regular pages in Microsoft Word™) was done by carefully designing the navigation for application. In order to work with the call for 'easy-to-find-information-while-walking-in-3-minutes' we have

- focused on the understanding of not only the screen of the PDA but also *the hardware as interface*. A short-

cut to the MINI-application was programmed and works by activation of a button at the front of the PDA.

- focused on gathering central information in *a MINI-frontpage* giving easy access to 1) instructions organized in archetypes with search function in the four areas presented in figure 3, 2) personal notes, 3) tables and 4) link to the Danish website for medical handbooks, drug catalogues etc. used very often – and carried in heavy books – by physicians. A click on the logo of Aalborg University always present in the top of the screen brings the user back to the MINI-frontpage (fig. 3a)
- focused on providing direct access to specific information via *search functions* facilitating free text search or providing access to information by means of lists of symptoms, diagnostics, treatments, and monitoring (fig. 3b)
- focused on providing overview of the rather long text for archetypes by giving a) an overview of headlines (designed in the editor and in co-operation with the archetype-writers/chief physicians) under which text can be unfolded by clicking the drop-down icon, b) giving 'at a glance' information on whether there are notes to the text by using a transparent icon of a document when no notes are available and making this icon clear when notes has been inserted. The same icon with the '+' is the icon used when inserting notes to the text (fig. 3c).

Fig. 3 - Screen-dumps from the MINI-application showing a) the frontpage of the application, b) search functions and c) a mini-mized digital archetype organized in headlines which are unfolded/folded by use of the drop-down icon and linked to notes by use of the note/paper icon.

A real challenge related to mobile interaction has not only been organizing of long texts but also mini-mization of rather large tables often used by physicians. Figure 4a shows an example of an often used table in 'real size' which, if minimized 1:1 would be impossible to read. An area in the table makes no sense if not related to other areas or at least to the columns and rows. Consequently we programmed the interaction in the table with fixed columns and rows and scroll within the areas (fig. 4b).

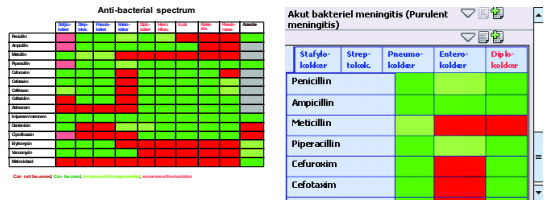


Figure 4 - a) a table in 'natural size', b) a screen-dump of the table in the MINI-application displaying the need for navigation here handled by keeping fixed columns and rows while scrolling in the table.

Ambitions of experimenting with voice interaction, primarily when making notes, failed due to technical problems but are on the top of a list for further research and development of the prototype.

Discussion: sustainability between user-drive and generic models

In the user-driven design process physicians were primary drivers behind the presentation and modeling of the clinical knowledge (their knowledge!) which was used in the design of the MINI-prototype by interaction designers. It is our impression that sustainable design cannot be identified at its source, in a (one!) practice (one information ecology) only. Likewise sustainable design cannot be validated on the generic potential of a model. Sustainability rather seems to lie in the interrelationship between the two – an interrelationship between user-drive and generic models which should not come as a surprise to software designers. However, it should emphasize (what cannot be emphasized enough) the importance of close collaboration between software developers and clinical personal in design processes. Further challenges and perspectives of the MINI-project are both the evaluation of the prototype but also inquiries into the use of PDAs at a national level. Both evaluations are taking place in the winter and spring of 2007.

Conclusions

The paper has presented results from the MINI-project on design of m-learning for physicians in hospital training. We have presented a) contextual findings on mobility and learning in the context of a medical ward, b) theoretical perspectives on medical knowledge, and c) description of our activation of this knowledge in the design of digital archetypes for PDAs.

It is argued that the user-driven innovation process of the project has provided a useful and valuable platform for sustainable development: it has both provided a platform for user-drive in the specific use-practice and provided the development of (steps towards) a generic model of clinical knowledge useful for further development and design projects. In other words, the user-driven innovation process has forced us to emphasise the combination of design and use and theory and practice – a combination which we have argued is essential to sustainable design.

Acknowledgments

We acknowledge the work of Pernille Bertelsen, Ellen Christiansen, and Thomas Christiansen and thank physicians at the ward for participating in observations, workshops, discussions, and pilot-test. The MINI-project is granted by the Danish Ministry of Science and Technology.

References

[1] Adolfsen, J. Problemer i videnskab, en erkendelsesteoretisk begrundelse for problemorientering (Problems in Science, an epistemological argumentation for problem-orientation), Aalborg: Aalborg Universitetsforlag, 1992

[2] Al-Ubaydli, M. Handheld computers. BMJ 2004: 328: 1181-1184

[3] Beale, T. Archetypes and the EHR. Stud Health Technol Inform 2003: 96: 238-44

[4] Bowen, J.L. Educational Strategies to Promote Clinical Diagnostic Reasoning. Medical Education 2006: 355: 2217-25.

[5] Dreyfus, H.L., Dreyfus S. Mind over Machine: the power of human intuition and expertise in the era of the computer. Oxford: Blackwell 1989.

[6] Fisher, S., Stewart, T.E., Mehta, S. Wax, R., Lapinsky, S.E. Handheld Computing in Medicine. Journal of the American Medical Informatics Association 2003: 10 (2): 139-49

[7] Kanstrup, A.M., Christiansen, E. Mobility at a Medical Ward: design challenges and decisions for an m-learning application. Proceedings of the Sixth Danish Human-Computer Interaction Resarch Symnposium (DHRS), Aarhus, Denmark, November 15 2006: 31-32

[8] Kanstrup, A.M., Christiansen, E. Selecting and Evoking Innovators: combining democracy and creativity. Proceedings of the 4th Nordic Conference on Human-Computer Interaction (NordiCHI), Oslo, Norway, October 14-18 2006: 321-330.

[9] Klass, D.J. Will e-learning improve clinical judgment? BMJ 2004: 328: 1147-1148.

[10] Lave, J., Wenger, E. Situated learning – Legitimate peripheral participation. Cambridge University Press 1991.

[11] Metcalf, D.S. mLearning – Mobile learning and performance in the palm of your hand. HRD Press, Inc. Amherst, Massachusetts, USA, 2006

[12] Nardi, B.A., O'Day, V.L. Information Ecologies – using technology with heart. MIT Press 1999.

[13] Rothschild, J.M. Fang, E. Liu, V., Litvak, I., Yoon, C, Bates, D.W. Use and Perceived Benefits of Handheld Computer-based Clinical References. Journal of the American Medical Informatics Association 2006: 13 (6): 619-626.

[14] Schön, D.A. The Reflective Practitioner – How professionals think in action. Basic Books 1983.

Address for correspondence

Anne Marie Kanstrup,
V-CHI/Dept. of Development and Planning,
Aalborg University,
Fibigerstraede 13, 9220 Aalborg, Denmark.
E-mail: amk@plan.aau.dk.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Evaluation and Assessment of the Online Postgraduate Critical Care Nursing Course

Reena Patel

School of Nursing, University of Auckland, New Zealand

Abstract

During challenging times facing the health service, strategies for sustaining further education for nurses in highly specialised areas call for alternate means for learning. Nurses, who were accustomed to traditional methods of learning and had no formal computer training as part of their curriculum, are now being confronted with new methods of learning. Evaluation of the effectiveness of a newly developed postgraduate critical care course delivered online for nurses was examined. A pre test and post test of 16 participants were conducted. Participants found coursework intellectually stimulating and their preference to learn from websites demonstrated the effect size (tau-b = .677) to be significant at the 0.01 level. The pre and post test results from the knowledge assessment tool indicated an advancement of mean test scores and at a significant difference value of p = .055. Ninety four percent of the participants agreed that they were able to integrate their learning from the coursework towards their clinical practice. Improvement in nurses critical care knowledge impacts positively on delivery of safe and effective health care.

Keywords:

education, nursing, critical care, computer assisted instruction

Introduction

Development of knowledge in the field of health is synonymous with using computer technology. The existence of the Internet and the vast amount of information available at the fingertips is the reality for nurses today. Computer technology is used to provide further education to nurses employed in critical care environments.

Traditional classroom settings are being replaced by the use of web-based and resource supplemented packages for continued learning [1]. To meet the needs for access and convenience, postgraduate nursing courses are being offered via distance learning environment [2]. As new web-based nursing courses emerge, the emphasis to examine quality means that evaluation and assessment need to be considered.

The increasing reliance on technology has become more apparent within the health science disciplines and has been integrated into many courses. The use of technology in nursing education has been directly extended into both graduate and postgraduate programmes. Nurses, who were accustomed to traditional methods of learning and had no formal computer training as part of their curriculum, are now being confronted with new methods of learning.

These new methods of learning for students encouraged more autonomy in learning [3]. Previously, the use of chalk and blackboards were the primary modes of communication in classes. Later, the use of overhead projectors, whiteboards and markers came to use. Eventually computers were integrated into the teaching and learning process. Today, the convenience of a learner having access to a course material anytime and anywhere [4] has removed the barriers of time and distance previously imposed by traditional classes [5]. The increased emphasis on technology in health care and nursing education should facilitate the development of flexible, knowledgeable and technologically competent practitioners [4].

This paper discusses the effectiveness of a newly developed critical care course delivered online as part of a postgraduate certificate program for nurses. Research is required to evaluate the effectiveness of online instruction, in the development of new knowledge and skills [6]. Additionally, the study examined the critical care nursing course for enhancing knowledge attainment and for student satisfaction with web based learning.

Methods

The longitudinal study design involved one study group of sixteen students. All participants completed the Online Post Graduate Specialty Practice Adult/Paediatric Critical Care nursing course. The design involved data collection at two intervals, prior course commencement, and at course completion. The questionnaire examined student application to course material and evaluation on the effectiveness of the online course. Critical care nursing knowledge was assessed utilising a tool known as the Australian Basic Intensive Care Knowledge Test (ABICKT) [7]. ABICKT was used to test participants prior to commencing the online critical care course. On completion of coursework at the end of semester two the test was reapplied to determine any difference in scores.

Results

Learning with technology

Results were obtained from a series of questions related to learning with technology. It was important to ask questions to the participants pertaining to whether additional learning and training with technology was required prior to or during the course. The participants responded to questions related to the ease in navigating around the course and their satisfaction towards the technical access to the course.

Prior to commencement of coursework, an introductory seminar regarding the means of accessing coursework and assessments was demonstrated. Participants were asked in the post course questionnaire if they had found that they required additional technological learning when course content was commenced. Of the 16 participants, 10 (62.5%) of them agreed that additional learning was required. Only 4 (25%) participants strongly disagreed that no additional technological learning was required when commencing the course content.

The post course questionnaire sought data on responses from participants regarding the ease of navigation around critical care course materials. Course materials consisted of Internet web site addresses, Microsoft PowerPoint presentations, online journal articles from the university course library page and coursework set out in Microsoft Word documents.

Results from this question indicated that 11 (68.8%) of the 16 participants found it easy to navigate around course materials and 5 (31.3%) found it somewhat difficult. Participants were then asked if they were satisfied with the technical access to the coursework. Of the 16 participants, 13 (81.3%) were satisfied with the technical access compared to 3 (18.8%) participants who disagreed to this statement.

Correlations between data obtained from participant's experiences with technology and levels of support were considered. The most significant correlation seen at the 0.01 level was between the data related to participant's satisfaction with technical access to coursework and their ease of navigation around the coursework material. The correlation was investigated using Kendall's tau_b correlation. Expectedly, there was a positive association between these two variables, p= 0.01 and tau_b = 0.55 (see Table 1).

Table 1 - Correlation of the variables for online learning experiences

		Ease to navigate in course	**Satisfaction of technical access for coursework**
Adequately prepared for technology used in course	Correlation Coefficient	.532(*)	.512(*)
	Sig. (1-tailed)	.012	.014
	n	16	16
Ease to navigate in course	Correlation Coefficient	-	.549(**)
	Sig. (1-tailed)	-	.009
	n	-	16

* Correlation is significant at the 0.05 level (1-tailed).
** Correlation is significant at the 0.01 level (1-tailed).

Interestingly, when the strength of the relationship between ease of navigation in the course and the participant being adequately prepared for the technology used in the course were considered, there was a strong correlation at the 0.05 level, p= 0.12 and tau_b = 0.53 (see Table 1).

Correlation between a participant's satisfaction with technical access to the coursework and being adequately prepared for the technology used was also considered. Again the results from Kendall's tau_b analysis showed a significant positive association between these two variables p= 0.01 and tau_b = 0.51, at the 0.01 level (see Table 1).

Coursework satisfaction

On completion of coursework and assignments, the post course questionnaire sought the responses from participants in relation to their advancement in critical care nursing knowledge. Of the 16 participants, 4 (25%) strongly agreed with this statement and 12 (75%) somewhat agreed (see Figure 1). There were no participants who disagreed with this statement. Further, they were asked if they were able to integrate course learning with their clinical practice. Of the 16 participants, 15 (94%) responded by strongly agreeing and somewhat agreeing to the statement (see Figure 1).

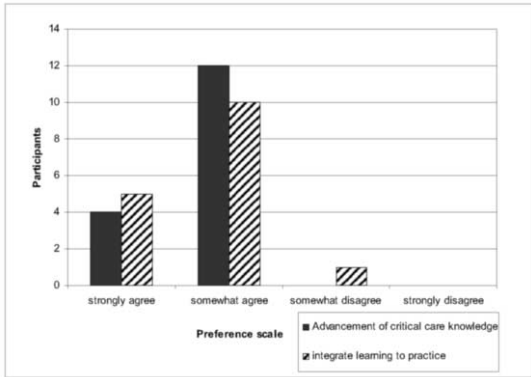


Figure 1 - Integration of advanced critical care knowledge to practice

The critical care course provided current world wide web (www) addresses as adjuncts to course learning. In the post course questionnaire participants were asked as to their preference for learning from these websites. Of the 16 participants, 11 (64.7%) preferred learning from websites and 5 (31.2%) indicated they did not prefer this method of learning (see Figure 2).

All participants were asked to indicate whether journal articles for each topic contributed to an increase in their understanding of the subject. Of 16 participants, 14 (87.6%) participants agreed that the journal readings deepened their understanding and 2 (12.5%) participants somewhat disagreed to this statement. There were no participants who strongly disagreed regarding the contribution of journal articles to the coursework learning and understanding (see Figure 2).

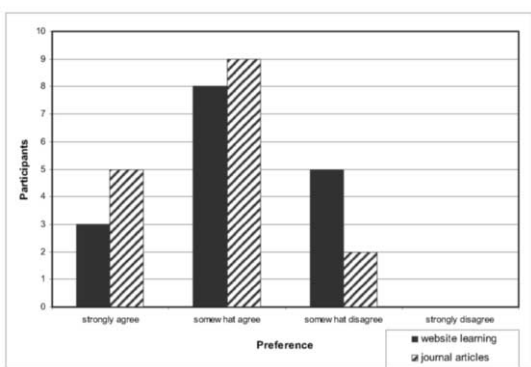


Figure 2 - Preferences of learning method

Importantly, there was a strong correlation between participants' use of online journal articles and their ability to integrate course learning to practice, p= 0.001 and tau_b = 0.726. Those participants who found coursework intellectually stimulating also reported a preference for learning from websites. This correlation was seen at a level of p= 0.002 and tau_b = 0.68, which was significant at the 0.01 level.

Critical care nursing knowledge test

Results from the pre course and post course critical care nursing knowledge test findings are presented. The post-test results revealed no changes in the number of higher scores (i.e. number of participants who achieved marks 80%). What it did reveal, however, was a positive shift of scores from the 60th percentile to the 70th percentile for several participants (see Figure 3).

The calculated median for the pre test results was 66, and the median for the post-test results was 71. This demonstrated an improvement of test scores post completion of coursework and assessments.

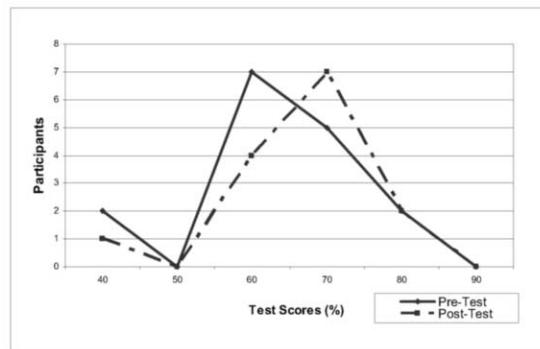


Figure 3 - Pre course and post course critical care knowledge test results

Wilcoxon Signed Rank test was used for analysis of the pre and post ABICKT results. It indicated that the two set of scores were approaching significant difference at p = 0.055. The median for the pre and post knowledge assessment test was 66 and 71, respectively. The median increased by 5% in the post-test. The majority of the participants achieved a score greater than 60% and below 90% in both the pre and post-test. There was a 10% increase in scores for some participants who scored 60% in the pre-test.

Discussion

The study examined the effectiveness of coursework learning and the application of the knowledge acquired to clinical practice. In nursing, the capacity to blend nursing education with clinical competency and problem solving skills has been a growing demand within the nursing profession. Dunn et al [8] and Morrison [9] have suggested that the ability to master the cognitive facts and apply the knowledge with critical problem solving and psychomotor skills in a clinical situation can be correlated with competency. The applicability of learning in the course to the clinical setting was relevant.

All participants in the course had indicated that they perceived their critical care learning had advanced. When questioned about how their learning was being applied to practice, 94% agreed that they were integrating their new knowledge to clinical settings. Nurses working in critical care particularly valued their training in physiology-based

decision-making. Saggs [10] highlighted this view and considered that physiology decision-making and practice skills were considered extremely relevant to postgraduate education in critical care. Equally significant, as discussed by Manias and Aitken [11], was the importance placed on nurses achieving technological competence in critical care.

In this investigation although there was a favourable outcome, one participant felt that their skills had not advanced in the clinical setting. In accordance to principles of Gestalt psychology, learning is perceived in different ways depending to the learner's experiences [12, 13]. The fact that this particular participant had not transferred new learning to clinical practice, suggests the level of exposure to clinical experiences need to be explored. Toth [14] concurred and proposed that increased knowledge occurs with more experience.

The relevance of content learned and its applicability to clinical practice was also examined in the study. Knowles et al [15] has pointed out that learning occurs not only when one is exposed to learning materials, but also occurs in the workplace. In the workplace, one is able to apply higher thinking processes leading to the interpretation and understanding of new learning. This has been well documented by learning theorists, who emphasised that adult learners need to know why they are learning something and then understand the meaning of it when applied to their real world [13, 16-18]. Content and intended learning should be related to the needs of the nurse in order for it to be applicable in practice.

Motivation to learn goes hand in hand with engaging the student with the learning material and ensuring the provisions of adequate intellectual stimulation. The results from this course evaluation showed that students perceived that the course had motivated them to learn, they had a clear idea of what was expected of them as they proceeded with coursework, which they were able to cope with. Online instruction engaged students in an interactive learning environment that extended beyond the classroom and increased their access to the wealth of educational resources online [6].

The online critical care course provided postgraduate study for nurses who were employed in critical care units. A variety of nurses were able to enrol into the course for example those who were new to practice, those seeking career development, or those who were re-entering into the critical care specialty. Students used technological competence to make sense of the critical care environment and to support and extend their existing practice. Wynd [19] proposed that critical care knowledge improved when nurses participated in specialised critical care courses which developed critical care experience. In this study, critical care nurses displayed an earnest desire to improve their knowledge and skills.

Interestingly, a strong association was seen between the variable of participant's use of journal articles for learning and the ability of participants to integrate course learning to clinical practice. A correlation of $p = .01$ and an effect size of tau_b - .726 was observed. Journal articles were all online and focussed on the topic objectives. Equally important was gaining information and learning from websites, yet only 65% of the participants preferred this mode of learning. This revealed differences in the manner students interacted with computer-mediated learning materials. Although learning involved printed material and computer mediated material, Benson [20] observed that learning via online instruction involved more than just obtaining lecture notes from technology mediated programs. It involved an integration of delivery options available for interactivity and learner engagement with the online content.

Interactivity has been considered to represent two-way pathway for the flow of information between the facilitator and the participant [1]. In this investigation, the high levels of interactivity have encouraged learners to participate in learning activities. As a consequence they have demonstrated enhanced retention, transfer of learning and satisfaction [20].

Conclusion

The use of technology in higher education has grown exponentially. In order to establish the worth and merit of these programs, careful evaluation needs to be conducted. Williams [21] has suggested that since there was no norm with which to compare how well learning with technology has worked, conclusions as to whether a program has in fact improved learning outcomes need to be examined critically.

The primary outcome of the research indicated that all participants perceived an increase of critical care nursing knowledge after completion of all coursework and assignments. Ninety four percent of the participants agreed that they were able to integrate their learning from the coursework towards their clinical practice.

The secondary outcome of the research indicated that the pre course and post course knowledge assessment test indicated an advancement of mean test scores and at a significant difference value of $p = .055$ in test scores.

As discussed by Williams [21], evaluation can prove to be a powerful partner for improving higher education, as long as relevant participants are involved in the process. It should be discussed in the context of its users and its purpose. To keep up with our changing society, nursing and nursing education should continue to develop innovative ways of teaching, reflecting the past but setting new directions for future knowledge and providing safe and effective health care.

Acknowledgements

This paper is based on research undertaken for Master of Health Science, Health Informatics, with thanks to Michelle Honey, fellow nurse and supervisor.

References

[1] Tegtmeyer K, Ibsen L, Goldstein B. Computer-assisted learning in critical care. Critical Care Medicine 2001; 29(8 suppl):N177-N182.

[2] Billings DM. A framework for assessing outcomes and practices in web-based courses in nursing. Journal of Nursing Education 2000;39(2):60-7.
[3] Jeffries PR. Development and testing of a hyperlearning model for design of an online critical care course. Journal of Nursing Education 2005;44(8):366-372.
[4] Magilvy JK, Smith MC. Chapter 9 Distance graduate education:The University of Colorado experience. In: Novotny J, editor. Distance Education in Nursing. New York: Springer Publishing Co; 2000. p. 136-152.
[5] Boyer SA. A comprehensive critical care course. Nursing Management 1996;27(6):32b-32f.
[6] Lashley M. Teaching health assessment in the virtual classroom. Journal of Nursing Education 2005;44(8): 348-350.
[7] Boyle M, Kenney C, Butcher R. Background information for users of the Australian basic intensive care knowledge test. Randwick, Australia: Prince of Wales Hospital; 1994.
[8] Dunn SV, Lawson D, Robertson S, Underwood M, Clark R, Valentine T, et al. The development of competency standards for specialist critical care nurses. Journal of Advanced Nursing 2000;31(2):339-346.
[9] Morrison S. Writing multiple-choice test items that promote and measure critical thinking. Journal of Nursing Education 2001;40(1):17-24.
[10] Saggs P. Developing postgraduate awards for critical care. Nurse Education Today 2003;23(4):307-314.
[11] Manias E, Aitken R. Achieving collaborative workplace learning in a university critical care course. Intensive and Critical Care Nursing 2003;19:50-61.
[12] O'Connor AB. Clinical instruction and evaluation: A teaching resource. Boston: Jones and Bartlett; 2001.
[13] Ally M. Chapter 1 Foundations of educational theory for online learning. In: Anderson T, Elloumi F, editors. Theory and Practice of Online Learning. Athabasca: Athabasca University; 2004. p. 3-31.
[14] Toth J. Basic knowledge assessment tool for critical care nursing version 4 (BKAT-4): Validity, reliability, and replication. Critical Care Nursing 1994;14(3):111-117.
[15] Knowles MS, Holton III EF, Swanson RA, editors. The Adult Learner. 5th ed. Houston,TX: Gulf; 1998.
[16] Knowles MS, Holton III EF, Swanson RA. The Adult Learner. 6th ed. London: Elsevier; 2005.
[17] Knox D. Learning Theories. In: McEwen M, Wills EM, editors. Theoretical Basis for Nursing. Philadelphia: Lippincott Williams & Wilkins; 2002. p. 321-344.
[18] Cutler LR. From ward-based critical care to educational curriculum 1: A literature review. Intensive and Critical Care Nursing 2002;18(3):162-170.
[19] Wynd C. Evidence-based education and the evaluation of a critical care course. The Journal of Continuing Education in Nursing 2002;33(3):119-125.
[20] Benson EP. Online learning: A means to enhance professional development. Critical Care Nurse 2004;24(1):60-62.
[21] Williams DD. Improving use of learning technologies in higher education through participant orientated evaluations. Educational Technology and Society 2002;5(3):11-17.

Address for correspondence

Reena Patel, School of Nursing, Faculty of Medical and Health Science, University of Auckland, Private Bag 92019, Auckland, New Zealand
reena.patel@auckland.ac.nz

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Development and Evaluation of a PDA-based Decision Support System for Pediatric Depression Screening

Ritamarie John[a], Penélope Buschman[a], Mosely Chaszar[a], Judy Honig[a], Eneida Mendonca[b], and Suzanne Bakken,[a,b]

[a] *Columbia University School of Nursing, New York, New York*
[b] *Department of Biomedical Informatics, Columbia University, New York, New York*

Abstract

Depression is under recognized in a variety of pediatric settings. The purpose of this paper is to describe the development and initial evaluation of a personal digital assistant (PDA)-based decision support system (DSS) for pediatric depression screening in ages 8 to 18 years of age by pediatric advanced practice nurse (APN) students. Three aspects are described: selection of depression screening instrument; integration of the instrument into the PDA; and quantitative (usage) and qualitative (focus group) evaluation. Only one third of eligible patients were screened. Twenty percent of those screened were identified as at risk for mood disorder. The barriers to screening identified through focus groups included a lack of time, knowledge, intervention protocol, referral resources, PDA usability issues, preceptor motivation and comfort, as well as perceived or real cultural barriers. Suggestions for educational, research, and interventions to integrate clinical based PDA-based screening are discussed.

Keywords:

depression, mood disorders, child, pediatrics, primary health care

Introduction

Childhood and adolescent depression is a serious public health problem [1]. Modern interest in child and adolescent depression arose in the late 1970 [2] and over the past 30 years, there has been an increased recognition of pediatric mood disorders [3]. It is estimated that depression occurs in 9 per 1000 preschoolers and 20 per 1000 school-age children [4]. The prevalence of a major depressive disorder (MDD) by late adolescence is ranges from 14 to 20% [5-8]. Depression is a risk factor for suicidal behavior [9], reduced social functioning [10], long term effects on neurocognitive functioning [11-12], poor school performance [13], serious conduct disorders [14], and drug use [15]. Untreated depressive disorders during childhood and adolescence can herald the onset of chronic and recurrent disorders during the person's lifetime [16-17]. Thus, early identification of children and adolescents at risk for a mood disorder is an important issue for clinicians.

While suicide can occur without previous depression in children and adolescents [18], depression is also sometimes a precursor to suicide During the last few decades, the suicide rate has steadily increased reaching a peak in the late 1990s [9]. Over the past few years the suicide rate has not changed substantially and at present, the rate of successful suicide is .6 per 100,000 in 5-14 year-olds and 9.7 per 100,000 in 15-19 year olds [19]. In addition, a history of prepubertal suicidal behavior predicts suicide attempts in adolescence [20-21]. Consequently, in screening for mood disorders, it is important to screen separately for suicidal ideation.

Health care providers in a variety of settings are in a unique position to identify, manage, and coordinate care for children with mental health disorders including depression [22]. Furthermore, the lack of child mental health specialists has placed an increased burden on primary providers to identify and treat children with behavioral health problems [23]. Multiple studies have confirmed that there is an under identification of mental health disorders in primary care settings [24-26]. If the provider does not use a standardized screening instrument, the detection of children with behavioral problems is lower [28-29].

At Columbia University School of Nursing, a PDA-DSS was designed to aid APN students to document and analyze patient encounters and to provide decision support for the screening and management of obesity, depression, and smoking cessation. A randomized controlled trial is in progress. The purpose of this paper is to describe the development and initial evaluation of the PDA-DSS for pediatric depression screening in ages 8 to 18 years of age by pediatric APN students. Three aspects are described: selection of depression screening instrument; integration of the instrument into the PDA; and quantitative (usage) and qualitative (focus group) evaluation.

Materials and methods

Selection of depression screening instrument

A Medline search was conducted combining the key words of mood disorders, pediatric depression screening, primary health care, and pediatric depression. In addition, an Internet search for pediatric depression screening was

conducted. The search yielded twenty different instruments with acceptable specificity and sensitivity statistics. Team members with expertise in pediatrics, psychiatry, and psychometrics evaluated the tools taking into account age range, length of questionnaire, and population that the instrument had been tested in.

Integration of the instrument into the PDA application

Questions from the screening instrument were added to the knowledge base that supports the PDA application. The content and algorithm were implemented for the Palm OS using the AppForge development environment.

Usage

SQL queries were developed to retrieve the data associated with depression screening from the central database. The following data were retrieved from the database for the time period of September 1, 2006 to November 26, 2006: the number of depression screening encounters versus the number of eligible depression screening encounters; medical diagnoses of those identified as at risk for a mood disorder and those not at risk for a mood disorder; characteristics of encounters in which pediatric depression screening occurred (patient demographics, family history of depression).

Focus groups

The objective of the two focus groups was to gain an understanding of the knowledge, attitudes, and beliefs of APN students regarding the use of a PDA-DSS. A purposive sample of 6 first and 6 fourth semester pediatric APN students were invited to participate in a 1-hour focus group. A 10-question, investigator-developed, interview guide was used to facilitate the discussion. The questions were developed from discussions with a panel of doctorally-prepared experts in informatics and qualitative research methods, as well as from a review of the qualitative literature on the attitudes of physicians [33-34] regarding PDA use in the clinical setting.

At the beginning of the group, each participant was given a copy of the interview guide. An experienced psychiatric clinical nurse specialist (PB) and an experienced pediatric APN (JCH) each conducted and facilitated one of the focus groups. An observer (RJ) took notes and audiotaped the discussion. During group discussion, the facilitator put forth a question and participants responded while interacting with one another. Probes were used to gain more detailed information and personal accounts. Data were content analyzed to extract major themes related to knowledge, attitudes, and beliefs.

Results

Depression screening instrument

After a careful review of the literature [30-32] and discussion among research team members, the team selected the Short Mood and Feeling Questionnaire (SMFQ) to measure risk for mood disorders rather than a longer diagnostic instrument. An addition four questions were added: two related to family history of depression [3] and two related to suicide [18].

Integration of the instrument into the PDA application

The SMFQ was displayed using 3-4 questions per screen (see screen shots below).

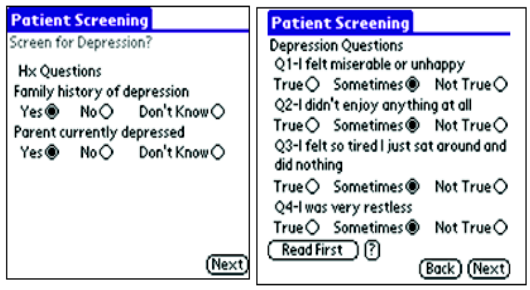


Figure 1 - Screen shots of application

The algorithm for the application is shown in Figure 2:

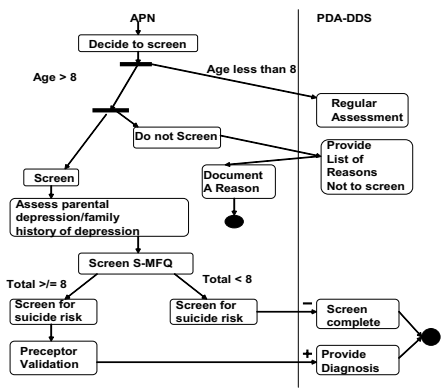


Figure 2 - Application algorithm

Usage

A total of 25 students were randomized into depression screening with 8 being fourth semester students and 17 being first semester student. In the fourth semester group, all students screened at least one child or adolescent for depression with an overall screening rate of 44.2%. The first semester students' overall screening rate was 37% with 23.5% doing no screening for depression. Students who failed to screen tended to be in specialty sites, ED follow-up clinic, or in private practices.

Of the 376 encounters eligible for pediatric depression screening, only one third (n=124) were screened for depression. If the student did not screen the child, they documented reasons for not screening. Over 81% cited other and about 8% noted a medical emergency as a reason for not screening. The remaining reasons included retardation (3.2%), learning disability (2.7%), patient refusal (2%), agitation (1.2%), other cognitive disorder (1.1%), and amnesic disorder (0.8%). Further explanation of the reasons for not screening in the other category is discussed in the qualitative data analysis.

Sociodemographic characteristics of patients in encounters in which screening versus no screening are shown in Table 1. Almost 80% of those screened were either Hispanic or Black.

Table 1 - Sociodemographic Characteristics in Screened vs. Not Screened (n=376)

	% Screened	% Not Screened
Ethnicity		
Hispanic	63.1	59.7
Black, not Hispanic origin	16.3	25.8
Asian or Pacific Islander	2.4	3.2
White, not of Hispanic origin	17.1	10.5
Other or unknown	1.1	0.8
American Indian or Alaska Native	0	0
Sex		
Female	51.6	52.4
Male	48.4	47.6
Age		
8-10	27.4	11.3
11-19	72.6	88.7

Overall, 20% of the patients screened were at risk for a mood disorder with 4% having a risk for a mood disorder and suicide. In the group at risk for a mood disorder, mean age was 13.1 years and 18.2% of patients were between 8-10 years. This was similar to the 13.4 years in the group at risk for suicide and a mood disorder. The latter group was totally female. No one screened was positive for suicide and negative for a mood disorder.

Within the group at risk for mood disorder, 35.3% had a family history of depression while 50% of those at risk for a mood disorder and suicide had a positive family history.

In terms of associated diagnoses, those identified as at risk of mood disorder had also had a high incidence of a psychiatric, behavioral or developmental disorders (n=44). The diagnoses with highest frequency were Attention Deficit Hyperactivity Disorder (ADHD), violence risk, and coping impairment. In the group at risk for a mood disorder, 68.2% had a behavioral or developmental disorder and only 31.8% had a diagnosis consistent with a physical disorder. However, of 467 diagnoses listed in the group not screened not screened for depression, only 6.6% had a psychiatric, behavioral, or developmental disorder diagnosis. The diagnosis with the highest frequency in the not screened group was well child (8.4%).

Focus group

Focus group discussions enabled exploration of students' knowledge, attitudes and beliefs about depression, depression screening, and the PDA-DSS. The major themes extracted for each of these areas are described in the following paragraphs.

Knowledge. Three main themes emerged during the focus groups: (1) limited, but some exposure to the problem of pediatric depression, (2) lack of a protocol within clinical site, and (3) lack of preceptor knowledge and support for depression screening. Many of the preceptors were not familiar or desirous of initiating screening.

Attitudes about depression screening. There were five related patterns centering on the theme of support that emerged during the focus groups. The students were concerned about a lack of time, knowledge, referral sources, comfort, and preceptor experience. The students expressed their need for more support from the site, preceptor, referral system, appointment system, and educational system to improve their knowledge base and comfort levels. The student felt it was inappropriate to do screening in emergency rooms or specialty clinics and wanted screening in primary care sites.

There was also a use and usability theme related explicitly to the PDA-DSS. Most students felt it interfered with the therapeutic relationship by creating a barrier between the patient and the student. One student commented, "The screening did not stimulate discussion. When I am done, there is not a transitional step." In terms of suicide screening, there was a similar theme of support. The students identified discomfort and apprehension regarding a positive screen. "You can ask the questions, but you must know what your next step is."

There were also cultural concerns raised since the students raised the issue of whether or not the patients felt comfortable with the idea of sharing their feeling. One student who was in a practice of recent Chinese immigrants commented, "Talking about feelings is not acceptable in this population."

Beliefs about barriers and benefits of depression screening. There was a discrepancy about the students' beliefs about depression screening and the barriers to depressions. Table 2 summarizes the themes identified.

Table 2 - Barriers and Benefits to PDA-Based Depression Screening

Benefits	*Barriers*
Prevent suicide *Enables sharing of feelings* *Opportunity to give holistic care* *Helps to pick up depression* *Improved quality of care*	*Time* *Perceived/real lack of referral resources* *Lack of preceptor knowledge and support* *Lack of knowledge of interventions* *PDA format* *Student discomfort with screening* *Cultural barriers*

Discussion

Clearly, there is a discrepancy between what the students believe are the benefits and their actual behaviors. Even though the students felt that screening should take place in a well visit, in 8.4% of the encounters in which no screening occurred, the diagnosis was a well child. In addition, only one third of eligible children or adolescents were screened. While recognizing that depression needed to be identified, the barriers overpowered the benefits. It is not clear whether the difference in the behavioral, develop-

mental, psychiatric diagnoses was a result of the greater awareness of the behavioral problems brought out by the depression screening or whether these children were presented with behavioral problems, cueing the students to screen.

This study has educational, clinical, and research implications. First, the students need to be empowered to intervene in children at risk for a mood disorder. Although teaching and counseling interventions are included in the PDA-DSS, students need further education about developing a therapeutic relationship and intervening with a single encounter. Educating preceptors about depression screening and interventions in a variety of settings may also promote screening.

The students need to improve their use of the PDA in clinical settings through strategies such as sharing the PDA screening with patients and entering initial information before entering the room.

Further research is needed to explore the patient's reactions to the PDA-based depression screening. It would also be important to compare different cultural group's reaction to the depression screening tools in paper versus PDA format. Further research is needed to assess whether the identification of depression helps students to identify other behavioral, developmental or psychiatric issues.

Conclusion

The data support that risk for mood disorder is fairly common in the population of primarily Hispanic and Black children and adolescents screened using the PDA-DSS and that there are missed opportunities for depression screening in a variety of settings. There are site, preceptor/student, and usability issues preventing depression screening. The organizational challenges are particularly difficult because APN students practice in different clinical sites with different resources and protocols. Since pediatric depression screening is relatively new, even experienced preceptors may lack the education to be effective primary contacts for the screening and identification of children at risk for a mood disorder or suicide.

Informatics innovations cannot be effectively implemented in practice without understanding the context of use. The triangulation of quantitative and qualitative methods was useful in helping the research team understand the screening behavior of students using the PDA-DSS and the associated organizational barriers in clinical settings. Moreover, these data provide direction for curricular changes and communication with preceptors regarding the use of the PDA-DSS.

Acknowledgments

This research is sponsored by a grant from the National Institute for Nursing Research (R01 NR008903). The authors thank the other members of the Mobile Decision Support for Advanced Practice Nursing for their advice and assistance. The authors also acknowledge the assistance of Dr. S. Ledlie in the design of the interview guide.

References

[1] Wells KB, Kataoka SH, Asarnow JR. Affective disorders in children and adolescents: addressing unmet need in primary care settings. Biol Psychiatry 2001;49(12):1111-20

[2] Pfeffer CR, Conte HR, Plutchick R, Jerrett I Suicidal behavior in latency age children: an empirical study. J Am Acad Child Adolesc Psychiatry 1979;18:670-692.

[3] Elliott GR Smiga S. Depression in the child and adolescent. Pediatr Clin N Am 2003;50:1093-1106.

[4] Besseghini VH. Depression and suicide in children and adolescents. Ann NY Acad Sce 1997;816:94-8.

[5] Lewinsohn PM, Rohde P, Seeley JR. Major depressive disorder in older adolescents: prevalence, risk factors and clinical implications. Clin Psychol Rev 1998;18(7):765-94.

[6] Demler O, Jin R Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS; National Comorbidity Survey Replication The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003;289:3095-105.

[7] Costello EJ, Mustillo S, Erkanli A, Angold. Prevalence and development of psychiatric disorders in childhood and adolescence. Arch Gen Psychiatry 2003;60(8):837-44

[8] Kessler RC, Berglund P, Demler O, Jim R, Merikangas KR, Walters EE. Lifetime prevalence and age of onset distributions of DSM-IV disorders in the National Comorbidity Survey Replications. Arch Gen Psychiatry 2005;62(6):593-602.

[9] Pompili M, Mancineli I, Girardi P, Ruberto A & Tatrelli r. Childhood suicide: a major issue in pediatric health care.Issues Compr Pediatr Nurs 2005;28(1):63-8.

[10] Lewinsohn PM, Rohde P, Seeley JR, Klein D N, Gotlib IH (2003) Psychosocial functioning of young adults who have experienced and recovered from major depressive disorder during adolescence. J Abnorm Psychol 2003;112:353-363.

[11] Arborelius L, Owens MJ, Plotsky PM, Nemeroff CB. The role of CRF in depression and anxiety disorders. J Endocrinol 1999;1600:1-12.

[12] Charmandari E, Kino T, Souvatzoglou E, Chrousos GP. Pediatric stress: hormonal mediators and human development. Horm Res 2003;59(4):161-79.

[13] Jellinek MS, Snyder JB. Depression and suicide in children and adolescents. Pediatr Rev 1998;19:255-64.

[14] Zalsman, G, Brent DA., Weersing VR. Depressive disorders in childhood and adolescence; an overview epidemiology, clinical manifestation and risk factors. Child Adolesc Psychiatric Clin N Am 2006;15:827-841.

[15] Greydanus DE, Patel DR. Substance abuse in adolescents: a complex conundrum for the clinician Pediatr Clin N Am 2003; 50:1179-1223.

[16] 2Lewinsohn PM, Rohde P, Seeley JR, Klein DN, Gotlib IH Psychosocial functioning of young adults who have experienced and recovered from major depressive disorder during adolescence. J Am Acad Child Adolesc Psychiatry 1999;38:56-63.

[17] Pine DS, Cohen P, Gurley D, Brook J, Ma Y. The risk for early adulthood anxiety and depressive disorders in adolescents with anxiety and depressive disorders 1998;34:566-578.

[18] Apter A King RA. Management of the epressed, suicidal child or adolescent. Child Adolesc Psychiatric Clin N Am 2006;15:999-1013.

[19] Eaton DK, Kann L, Kinchen S, Ross J, Hawkins J, Harris WA, Lowry, R, McManus R, Chyen D, Shanklin S, Lim C, Grunbaum J,Wechsler H. Youth risk behavior surveillance United States, 2005. MMWR. 2006 Jun 15; / 55(SS05):1-

108 Available from http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5505a1.htm
[20] Pfeffer CR, Klerman GL, Hurt SW, Leser M, Peskin JR & Siefker CA. Suicidal children grow up: Demographic and clinical risk factors for adolescent suicide attempts. J Am Acad Child Adolesc Psychiatry 1991;30:606-616.
[21] Pfeffer CR, Klerman GL, Hurt SW, Kakuma T, Peskin JR, Siefker CA Suicidal children grow up: rates and psychosocial risk factors for suicide attempts during follow-up. J Am Acad Child Adolesc Psychiatry 1993;32:106-113.
[22] Kelleher KJ, Campo JV, Gardner WP. Management of pediatric mental disorders in primary care; where are we now and where are we going?
Curr Opin Pediatr 2006;18:649-653.
[23] Weitzman CC, Leventhal JM. Screening for behavioral health problems in primary care.
Curr Opin Pediatr 2006; 18:641-648.
[24] Williams J, Klinepeter K, Palmes G, Pulley A, Foy, JM. Diagnosis and treatment of behavioral health disorders in pediatric practice. Pediatrics. 2004; 114:601-606.
[25] Sayal K, Taylor E. Detection of child mental health disorders by general practitioners.
Br J Gen Pract 2004;54:348-352
[26] Brugman E, Reijneveld SA Verhulst f, Verloove-Vanhorick SP. Identification and management of psychosocial problems by preventive child healthcare. Arch Pediatr Adolesc Med 2001;155:462-469.
[27] Reijneveld SA, Brugman E, Verhulst F, Verloove-Vanhorick SP. Identification and management of psychosocial problems among toddlers in Dutch preventive childhealthcare. Arch Pediatr Adolesc Med 2004; 158:811-817.
[28] Simonian S, Tamowsi K. Utility of the Pediatric symptom checklist for behavioral screening of disadvaged children. Child Psychiatry Hum Dev 2001;31:269-278.
[29] Haughey MT, Claderon Y, Torres S., Nazario, S, Bijur P, Identification of depression in an inner city population using a simple screen. Acad Emerg Med 2005;12(12):1221-1226.
[30] Dahlberg LL, Toal SB, Swahn M Behrens CB. Measuring violence-rated attitudes, behaviors, and influences among youths: a compendium of assessment tools, 2nd ed., Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control, 2005. Available from http://www.cdc.gov/ncipc/pub-res/pdf/YV/YV_Compendium.pdf
[31] Myer K, Winters N Ten-year review of rating scales. II: Scales for internalizing disorders. J Am Acad Child Adolesc Psychiatry. 2002;41(6):634-659.
[32] Angold A, Costello E, Messer S. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. Int J Meth Psychiatric Res 1995;5:237-249.
[33] McAlearney AS, Schweikhart SB, Medow MA. Doctor's experience with handheld computers in clinical practice: qualitative study. BMJ. 2004; May 15;328(7449):1162
[34] McAlearney AS, Schweikhart SB, Medow MA. Organizational and physician perspectives about facilitating handheld computer use in clinical practice: results of a cross-site qualitative study. J Am Med Inform Assoc2005;12:568-575.

Address for correspondence

Ritamarie John, DrNP, CPNP
Columbia University
Mailbox 6
New York, NY 10032
Telephone: 1-212-305-5542
Fax: 1-212-305-6937
e-mail: rmj4@columbia.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Data Mining Results from an Electronic Clinical Log for Nurse Practitioner Students

Patricia Trangenstein[a], Elizabeth Weiner[a], Jeffry Gordon[a], Ryan McNew[a]

[a] *Frist Nursing Informatics Center, School of Nursing, Vanderbilt University, Nashville, TN, USA*

Abstract

Traditional techniques for collecting data on clinical experiences have been greatly flawed. Data cannot be easily collected in real time to make programmatic or placement changes "on the fly". Furthermore, it is difficult to look at data across students, specialty areas, and years because the data is typically not in a digital format. In response to this problem, the Vanderbilt University School of Nursing has created a web/PDA based clinical log to document the kinds of clinical experiences the students are having. Since our initial report, three years ago, we have collected three years worth of data, over 220,000 different patient encounters. This past year the data has been very complete, giving a full picture of the types of experiences the students are having. Our faculty have begun to analyze the data in the clinical log to examine the kind of experiences the students are having and to make programmatic changes and placement adjustments in real time. In general, the results supported that students in the various specialties managed patients and performed services appropriate to their specialty. Patients varied in ages, ethnic groups, payment sources, and medical diagnoses. Students did progress from an observer role to a more independent role in either a linear fashion or in a more biphasic mode with an increase in the observer role at the start of a new semester

Keywords:

education, nursing, graduate; nurse practitioners professional practice; students, nursing; computers, handheld/utilization; education, nursing, graduate/ methods; nursing education research; preceptorship

Introduction

Traditional clinical education of nurse practitioner (NP) students has used preceptors for direct clinical supervision with faculty responsible for student progression, student evaluation and attainment of course objectives as well as curriculum revision. Nursing faculty collected data from nurse practitioner students to document their clinical experiences via on-site evaluation, discussions with the preceptor, paper logs and scan sheets completed at selected intervals during the course of study. These models led to a significant delay in receiving the information from students and providing appropriate feedback. In addition, it was very difficult to aggregate this data for a given student, or across students for a given semester or program because of the non-digital nature of the media Furthermore, it was difficult for students to examine their practice over time and for faculty to create reports and track student's progress.

A real-time method of data collection and analysis would allow for changes during, rather than after, the semester of clinical practicum. Faculty also thought that it was appropriate to map current evaluation methods against competencies being developed by the various specialty groups as well as the National Organization of Nurse Practitioner Faculty (NONPF). In short, the faculty wanted to re-design their clinical evaluation tool for the students in their specialty while preserving the ability to make comparisons across specialties in an aggregate format for purposes of accreditation.

A review of the literature revealed a scarcity of information with only one article describing the content and quality of a precepted clinical experience for nurse practitioner students. [1] As a result, the informatics team was invited to work with the specialty directors in the design, implementation, and evaluation of an online web based clinical log with PDA data input capabilities.

Methods

In 2004 VUSN began implementing an electronic (web/ PDA based) clinical log for nurse practitioner students. By 2006, the clinical log had been used by 200 students per year in seven different nurse practitioner programs for the entire length of their clinical experience. This article presents the results from the analysis of over 89,000 records collected during this past academic year. The implications for nurse practitioner education, curriculum revision, and student clinical portfolios for future employment are substantial. Not only was the log data used for program evaluation through data mining techniques, the students also used the log to create an electronic portfolio of their experiences and have, in some cases, secured employment as a direct result of their log entries.

Description of the system

The electronic clinical log was initially designed for data entry through the web. However, students raised the issue that they were entering data twice (once on paper patient side and again that evening as they transferred the data to

the log) a PDA interface was designed in response to their concerns. Students were instructed to synch their PDA with their desktops every night to avoid data loss. They were further instructed to upload their data once each week to the server based database. The records in the database were accessible via a web browser by students, faculty, and preceptors. The students could see only their own data, faculty and preceptors could see their students' data. Data collected included patient's age, gender and ethnicity; type of insurance, services rendered (professionally recognized criteria designed by the National Organization of Nurse Practitioner Faculty (NONPF)), ICD9 or DSM-IV codes and self assessment of the student's responsibility in patient management. Students and faculty could view the records and export logs to a spreadsheet for aggregation and graphing. Because of the unique nature of the Psychiatric/Mental Health Nurse Practitioner program, the PDA component for this specialty will be discontinued. Instead, next year, the PDA will be used to actually audio record the patient encounter, uploaded to the log, and made available for listening and comment by the faculty member or preceptor.

All electronic clinical logs (N = 114,206) collected during the entire year of study were subjected to data cleaning prior to analysis. If a student did not have data for the entire program of study all their records were eliminated in this analysis. After data cleaning a total of 89,401client encounters for 200 students remained for the purpose of this study. SPSS was used for all an analyses.

Results

Each time a student interacted with a patient it was treated as a single encounter. Therefore, if the student saw the same patient multiple times, as is common in the Psychiatric/Mental Health Nurse Practitioner specialty, each visit counted as one encounter.

The total number of patient encounters per nurse practitioner program was astounding and ranged from 2,072 for Pediatric Acute Care Nurse Practitioner students (PNPAC) to 34,700 for Family Nurse Practitioner (FNP). On average, the data showed there were 10 cases per week for each Acute Care Nurse Practitioner student and 21.4 cases per week for each Family Nurse Practitioner student. The mean number of encounters per student ranged from a low of 250 for the entire year for Acute Care Nurse Practitioner students to a high of 550 for Nurse Midwifery students. It should be noted that most of the specialties have an 8 month clinical rotation; however, the Nurse Midwifery program had a 13 month clinical experience. In general, Acute Care Nurse Practitioner, Pediatric Acute Care Nurse Practitioner and Psychiatric/Mental Health Nurse Practitioner students had fewer patients but longer encounters.

Acute Care Nurse Practitioner students and Adult Nurse Practitioner students (ANP) students saw older patients with a mean age of 55 while students in the Family Nurse Practitioner (FNP), Nurse Midwifery and Psychiatric/Mental Health Nurse Practitioner (PMHNP) programs saw middle aged patients (average age = mid 20's to early 30's). Students in all programs saw an almost equal percentage of males versus females except for Family Nurse Practitioner (FNP) and Psychiatric/Mental Health Nurse Practitioner (PMHNP) students who saw almost twice as many females as males, and of course the Nurse Midwifery students saw only female patients.

While the nurse practitioner students in all programs cared for many ethnic groups, the majority of the patients were Caucasian (49 – 84 percent across all programs), African Americans (11-24 percent across all programs) or Hispanic (1.5 – 24 percent across all programs). The majority of patient seen by Adult Nurse Practitioner (ANP) and Acute Care Nurse Practitioner (ACNP) students used either Medicare or private insurance as their source of payment, while the majority of patients seen by Family Nurse Practitioner (FNP), Nurse Midwifery, Pediatric Nurse Practitioner (PNP) and Pediatric Acute Care Nurse Practitioner (PNPAC) students had private insurance or TennCare (replacement for the state of Tennessee's Medicaid program).

The log allowed the students to enter up to four ICD9 codes per record. DSM codes were supported for the PMHNP students. Tables 1-3 describe the top ranking ICD9 or DSM-IV codes for student encounters across programs. While there were some similarities, the data revealed that students were caring for patients that were appropriate for their specialty. For example, Acute Care Nurse Practitioner (ACNP) students managed more CHF patients, and a significant portion of the Family Nurse Practitioner (FNP) students encounters were routine infant or child health checks.

Table 1 - Rank order of ICD-9 Codes for ACNP, ANP and FNP

ICD-9	ANP	ACNP	FNP
250, Diabetes	1	2	3
401, Essential HTN	2	5	
401.1, Benign essential HTN	3	1	1
272.4, Other hyperlipidemia,	4	3	
272, Disorders of lipoid metabolism	5		
428, Heart failure,		4	
401.9, Unspecified essential HTN			2
V20.2, Rout infant or child health check			4
477.9, 461.9 Allergic rhinitis or sinusitis, 465.9, Acute URI			5

Table 2 - Rank order of ICD-9 Codes for NMW and PNP

ICD-9 Codes	NMW	PNP
V22.1, Supervision of other normal pregnancy	1	
V22.2, Pregnant state, incidental	2	
V20.2, Routine infant or child health check		1
465.9, Acute upper respiratory infections of unspecified site		2
382.9, Unspecified otitis media		3
460, Acute nasopharyngitis [common cold]		4

Table 3: Rank order of DSM-IV codes for PMHNP students

DSM-IV Codes	Rank Order
296.9Mood disorder	1
309.81PTSD	2
314.01 ADHD combined subtype	3.5
296.33 Major depression without Psychotic features	3.5
296.8Bipolar	5
304.8 Polysubstance Dependence	6

The faculty members were also interested in the types of services their students provided to their patients. The services rendered were derived from the NONPF content competencies and were customized by specialty. Only those programs with some commonality with other specialties were reported here.

Acute Care students (ACNP and PNPAC) provided an average of 10 services per encounter while primary care student provided an average of 3 services per encounter. Tables 4 and 5 describe the highest ranking services provided as a percentage of the encounters. For example, Acute Care Nurse Practitioner

(ACNP) students identified and documented actual patient problems almost 72% of the time while Pediatric Acute Care Nurse Practitioner (PNPAC) students provided patient and family education nearly 80% of the time.

Table 4 – Types of service rendered as a percentage of encounters for acute care specialties

Services Rendered	ACNP	PNPAC
ID & documented actual pt prob	71.9	68.3
Analyzed all pharmacological agents	68.9	53.1
Collaborated with others	65.4	69.6
Obtained a H&P	63.8	58.9
ID & documented potential pt prob	57.7	63.9
Provided pt and family education	56.0	79.5
Developed an individualized Rx plan	54.4	69.5
Documented H&P	52.5	58.5
ID expected outcomes ind'l to the pt	51.2	59.2
Developed & documented list of differential diagnoses	48.5	57.6
Utilized evidence based practice	47.7	52.4
Documented ongoing eval of Rx plan		55.2

Table 5 – Types of service rendered as a percentage of encounters for primary care specialties

Services Rendered	ANP	FNP	PNP
Focused/episodic exam	84.6	85.7	68.3
Prescription	66.0	73.3	37.7
Health Education	47.6	71.9	55.7
Labs	45.0	35.6	18.5
Complete H&P			34.1
Developmental screening			25.2

The faculty members also wanted to evaluate the level of a student's responsibility in patient management, increasing from observational roles in early clinical experiences to more independent roles later in their clinical experiences. The student responsibility ranged from Observer, to Novice, to Beginner, to Advanced Beginner and had descriptors associated with each level for reliability purposes. Most of the students' clinical experiences transpired over two semesters. In order to determine if there was a natural progression from observer to more independent roles, the date of the encounter was converted to the month of the year indicating a progression in clinical experiences.

Correlations between students' responsibilities and the month were statistically significant for all of the programs at the .001 level. The correlations ranged from .192 for the Pediatric Nurse Practitioner (PNP) students to .396 for the Acute Care Nurse Practitioner (ACNP) students indicating that students did indeed progress from observer to advanced beginner during their clinical experiences.

While the correlations were statistically significant, they were low, indicating significant variability. In order to visualize where the variability existed histograms were constructed for each of the specialties. Figures 1 and 2 indicate two different paths from observer to advanced beginner that were noted.

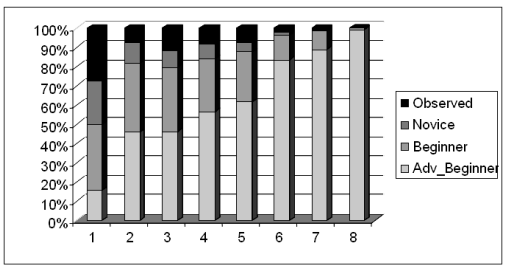


Figure 1 - Histogram of FNP students' responsibility by month of clinical experience

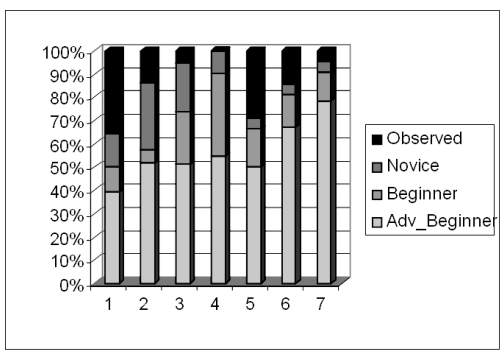


Figure 2 - Histogram of PNPAC students' responsibility by month of clinical experience

As can be seen from these two figures, the Family Nurse Practitioner (FNP) students showed a more orderly progression from observer to advanced beginner. On the other hand, the Pediatric Acute Care Nurse Practitioner students (PNPAC) students showed a more biphasic progression with an increase in the observer role in the fifth month. This corresponded to the beginning of a new semester and may reflect a change in preceptors, sites or performance expectations with the introduction of new skills. In any case, the faculty of the Pediatric Acute Care Nurse Practitioner students (PNPAC) program will look at the data to see if this trend continues and needs to be addressed or if it was an anomaly.

Discussion

NP faculty members have the ultimate responsibility for the supervision and evaluation of nurse practitioner students and for oversight of the clinical learning environment. [2] It is their responsibility to develop, evaluate and revise NP curricula. Students wish to secure the most favorable position possible and wish to make their case for such placement using data and evidence collect through the log. The NP faculty at the Vanderbilt University School of Nursing (VUSN) in conjunction with members of the Frist Nursing Informatics Center and with significant student input, developed electronic clinical logs (ECLS) for seven different nurse practitioner programs. Data presented represent a full program of clinical experiences for the students.

In general, the results supported that students in the various specialties managed patients and performed services appropriate to their specialty. Patients varied in ages, ethnic groups, payment sources, and medical diagnoses. Students did progress from an observer role to a more independent role in either a linear fashion or in a more biphasic mode with an increase in the observer role at the start of a new semester.

Conclusions

This article demonstrated the effectiveness of an electronic clinical log (ECL) for documenting the content and quality of a precepted clinical experience for NP students. Students entered data either by uploading from their PDA or by completing a web-based form. The records in the database were accessible via the internet by students, faculty, and preceptors. Data collected included patient's age, gender and ethnicity; type of insurance, services rendered (professionally recognized criteria designed by the National Organization of Nurse Practitioner Faculty (NONPF), ICD9 or DSM-IV codes and self assessment of the student's responsibility in patient management. Students and faculty could view the records and export logs to a spreadsheet for aggregation and graphing. Using informatics tools over 89,000 patient encounters for 200 NP students were analyzed.

Proposed revisions of the electronic clinical log (ECL) would include the creation of report dashboards to allow faculty members to determine "on the fly" whether the clinical experience was providing the right environment for the student. This would allow the faculty member to discuss with a preceptor the type of experiences the student was participating in as well as the quality of work the student was demonstrating. Currently a running total of the number of clinical hours has been well received by both students and faculty as a first piece of a dashboard. Given the volume of patient encounters the entire dashboard process would need to be automated with inadequacy flags automatically sent to the faculty member. We also plan on exploring the ICD9 data in detail to generate a culled-down list that would be more manageable for the students. We believe this shorter list will be far more useful to the

students who won't have to query through over 14,000 ICD-9 different codes.

Students have used the data to create a "clinical portfolio" listing their skills and activities across all of their clinical practice sites in their educational program and have analyzed their data using Excel. These advanced practice nurses will be participating in research and data collection/ analysis throughout their careers. A major benefit of our approach was that these students learned how to use Excel and demonstrated spreadsheet analysis techniques. The faculty were pleased that their students were as proficient with Excel as they are with Word.

An ECL will allow faculty to individualize learning based on the identification of gaps in the student's clinical experience. Faculty members can track the student's progress and types of patients seen. Data from the ECL can also be used to modify the curricula and provide documentation for grants and accreditation.

References

[1] Longworth, JCD, and Lesh, D. Development of a student nurse practitioner computerized clinical log. J. of the American Academy of Nurse Practitioners 2000:12 (4): 117-122.

[2] National Organization of Nurse Practitioner Faculties. Criteria for the evaluation of nurse practitioner programs: A report of the national task force on quality nurse practitioner education. National Organization of Nurse Practitioner Faculties, 2002.

Address for correspondence

Trish.Trangenstein@Vanderbilt.Edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Phase I Implementation of an Academic Medical Record for Integrating Information Management Competencies into a Nursing Curriculum

Carole A. Gassert, PhD, RN[a], Katherine A. Sward, MS, RN[a]

[a] *University of Utah School of Nursing, Salt Lake City, UT, U.S.A.*

Abstract

This paper is the report of the first phase of a case study from the University of Utah to help students and faculty integrate electronic information management into the nursing curriculum. Cerner AES, a live-production clinical information system with an academic overlay, has been implemented into the first semester of an undergraduate nursing program. A consortium of schools that use Cerner AES collaborate in the design and implementation of forms used by students. The consortium also allows members to share strategies for using the system. By using the system students are developing needed informatics competencies for beginning level nurses. The paper discusses the implementation strategies used and initial results of this project. Plans for expanding the project throughout the nursing curriculum are also presented.

Keywords:

electronic health record, nursing education, nursing informatics competencies, clinical information system, simulation

Introduction

Sustainable development can be defined as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" [1]. This concept has been adapted in the context of health care to define the capacity of a health system to continue over time without major interruptions [2].

Basic resources of sustainable health systems include data and knowledge, and the skills of the health workforce. Improving the safety, quality, and efficiency of health care systems requires ubiquitous access to complete patient information and decision support—that is, electronic health systems [3].

Information has become a capital good [4]. Nurses and other health professionals need skills that will allow them to more effectively manage information technologies. These skills must be learned from the beginning, during professional education [3, 5]. Staggers, Gassert & Curran [6, 7] identified information management competencies for nurses at multiple levels of practice. Despite recommendations by federal agencies that informatics content needs to be included in nursing curricula, information and information technology competencies have been slow to become part of the nursing curriculum [8] and many nurses may not be adequately prepared to manage information using technology [4, 9, 10]. A recent study of university students found that students' self-reports of information management activities were not an accurate predictor of their actual health information competencies [11]. Information management competencies should be assessed by observing actual use of an information system. Education regarding health information management should be conducted in the context of "real-world" applications and behaviors; that is, in the information environment where clinicians work, and should incorporate concepts such as confidentiality, systems thinking, and knowledge-resource evaluation [12].

One initiative that aims to increase the use of technology in nursing is the TIGER (Technology Informatics Guiding Education Reform) initiative [13]. This initiative, spearheaded by nursing and informatics leaders in the United States, has developed a 3-year plan to integrate informatics seamlessly into nursing.

Materials and methods

This paper describes the first phase of a case report at the University of Utah to incorporate information management competencies throughout the nursing curriculum by using a live-production application that simulates a clinical information system titled UCARE AES. The acronym stands for Utah Clinical Academic Record Excellence and is the name of the Cerner Academic Education Solution system installed for use in first semester undergraduate courses in May 2006. This on-going project is used to introduce curricular changes necessary to ensure that nursing students are adequately prepared to contribute to sustainable health systems. UCARE AES enables faculty to address many of the informatics competencies needed for beginning nurses identified by Staggers, Gassert, and Curran in their work. Examples of beginning level competencies that are taught in first semester undergraduate courses using UCARE AES are presented in Table 1.

Table 1 - Examples of beginning level competencies obtained through UCARE AES

Competency	Activity Example
Uses administrative applications for practice management	Searches for patient Retrieves demographics
Uses sources of data that relate to practice and care	Locate "patient data" in UCARE AES
Accesses, enters, and retrieves data for patient care	Charts on class activities in UCARE AES Creates plans of care in UCARE AES
Uses an application to document patient care	Uses UCARE AES to document results of class activities and simulated patients
Uses an application to plan care for patients to include discharge planning	Uses UCARE AES to document plan of care
Uses networks to navigate systems	Uses UCARE AES (hosted remotely)
Describes patients' rights as they pertain to computerized information management	Learns HIPAA regulations through UCARE AES
Identifies basic components of the current information system	Knows components of UCARE AES

Cerner academic education solution

An academic version of an electronic medical record is available through Cerner Academic Education Solution (AES), an application that simulates a clinical information system using Cerner's PowerChart. AES has an academic overlay that provides students with prompts and evidence-based practice information as they learn to document assessments and simulated patient events.

AES was first introduced at Kansas University [14, 15] and later at University of Missouri-Kansas City. With the introduction of UCARE AES at University of Utah, the three schools have formed a consortium to share design and management ideas and issues about AES. The consortium is a great forum for exploring new ways to use AES. Although the three schools share a domain server and databases that are hosted at Cerner headquarters in Kansas City, each school has a unique username and established structure. Many forms are shared as each school contributes their wisdom to the AES development process. All three schools are evaluating student outcomes. Monthly virtual meetings, held to discuss ideas and issues, are chaired by Cerner personnel.

UCARE AES implementation

It is interesting that each of the schools in the consortium has implemented the AES system differently. At the University of Utah planning for the phased implementation began in January 2006. Undergraduate faculty responsible for teaching the first level concepts of nursing, patient assessment and clinical care courses met, in spite of weekly snow storms, with the undergraduate program director and two informatics faculty in charge of the Cerner project to form an implementation team. The weekly meetings were held in a room that allowed access to and visualization of Cerner AES with an overhead projector. Minutes recorded decisions made by the team for later reference. The first task was for the entire implementation team to learn about AES and plan the structure for Utah. The team selected the name UCARE AES for the system and Swoopes Medical Center for our username. Units in the medical center were named for the focus of the semester, e.g., medical-surgical 1, medical-surgical 2, pediatrics, maternity, etc. and are created to accommodate the number of clinical groups assigned to the semester. For example, there are nine clinical groups in first semester so units Medical-surgical 1A through Medical-surgical 1I were developed. Students are then admitted as patients to their assigned "unit" and given a password to access the system.

To help students and faculty become more familiar with UCARE AES a number of "play patients" were admitted to the Utah Start Unit. Nicknamed the "Olympians" because their fictitious names reflect Olympic events, UCARE AES users were encouraged to chart anything on these patients, whether it made sense medically or not. Case studies for student use during the semester were created and admitted to a folder titled "UCARE 1." These patients were "off limits" for charting. When the semester began students located information needed for their assignments from the case studies, they charted the information requested on themselves as patients. To help students learn the principles of data security and confidentiality, they are warned that faculty could track where they have been on the system and if they are looking at fellow students charting.

Forms needed by students to complete their assignments were moved from the large repository of forms maintained by Cerner to specific folders designated by faculty in the first semester. They chose to have 5 folders – assessment, patient care, plan of care, clinical prep, and competencies. Most activities required forms from the first three folders. To make it easier for students to find needed forms, the Course Coordinator for the clinical course in first semester included instructions in her course syllabus. Examples are listed here:

Document the following on UCARE AES

- Vital signs: Located in Assessment Folder/Adult Vital Signs
- Results of incentive spirometer use: Located in Assessment Folder/Adult Respiratory Assessment/Incentive Spirometer (Left margin)
- Pulse oximeter results: Located in Patient Care Folder/Adult Vital Signs
- Teaching patient to TCDB: Located in Patient Care Folder/Adult Vital Sign/Oxygen Therapy/Document under "other."

- Use of O2: Located in Patient Care Folder/Adult Vital Signs

To allow the system built to be completed by Cerner in a timely manner, all design development was completed by mid-April for the May 15th go-live. First semester students and faculty were oriented to UCARE AES on the go-live day. Base-line data for first semester students' experience with and knowledge about information systems was collected at their orientation.

The system was implemented with accelerated baccalaureate students (students who hold a baccalaureate degree in another field and wish to become nurses) on their first day of clinical classes. Implementation team members and Cerner representatives were available during the first two days of clinical classes to help students sign onto the system and chart information required for class. The only problems encountered during the go-live were related to student IDs entered into the system, all problems were quickly resolved. Assistance was available in the Computer Lab to work with faculty and students as needed. During the planning phase, some faculty became champions for UCARE AES and were an essential part of the successful implementation of the system. It is also important to note that we had the full support of the Dean of the College of Nursing for implementing UCARE AES.

Since the implementation team is quite small, a decision was made to implement UCARE AES one semester at a time. As soon as the go-live was completed, second semester faculty responsible for clinical courses were asked to join the implementation team to plan for Phase II of the implementation. The plan is to have all four undergraduate semesters implemented by May 2007. Therefore, UCARE AES will have been implemented in the entire undergraduate curriculum in 13-14 months, an ambitious plan.

Results

Phase I student and faculty response

During the planning for implementing Phase I of UCARE AES, first semester faculty decided to tie the learning activities of the three courses more closely together. Timing of content was adjusted so students were learning the same concepts in all three classes. Faculty also began to standardize language used to teach concepts to students. During the design work, only one new form was created and one form was revised. Consortium members adopted the form that was developed for Utah. It seems that schools can use forms designed for others and such action will standardize some of the language used to teach concepts to undergraduate students across schools of nursing.

Data were collected at the beginning and end of the students' first semester. Data were also collected from the faculty at the end of the first semester. Specific data will be reported as aggregated data when more phases of the project have been completed. In general, faculty responses were mixed after the first semester of use; faculty reported feeling only moderately comfortable with how to use the UCARE system. Some faculty did not appear to grasp the purpose of learning to use an electronic record; with complaints about documenting on forms instead of free-text notes, and did not recognize the importance of teaching information management concepts, as reflected in comments such as "UCARE does not match the charting used at my clinical site". Other faculty recognized UCARE as a tool to teach information management, and commented that it was "easy to learn" and "an effective way to teach and learn terminology."

Our observations of nursing students concur with findings that student self-report of skills is likely not accurate. Incoming students rate their technology skills and knowledge as being low, and indicate they do not know how to find and manage information in an electronic health record, yet rate their information management skills as moderate. After one semester of use, the students reported higher technology skills and knowledge, and greater ability to find and manage information in an electronic record. However, these same first semester students also indicate that they want more guidance and direction with using the system.

The students were aware that their faculty were not yet comfortable with the system. Student comments were mixed, ranging from "busy work", "I don't get the point" and "it did not match the system at my clinical site"; to comments such as "Even though my clinical site has a different setup for their electronic records it still helped me" and "It was difficult at first to use not because of the system but because of my lack of nursing knowledge"; to comments such as "It was very cool to know about before starting clinical and to be able to say I know what to do - and how to get on and navigate the system."

Implementation status

The UCARE AES is an ongoing project that will eventually include undergraduate and graduate nursing, medicine and pharmacy students. Nursing is responsible for the initial design and has begun implementation at the undergraduate level. We recognize that repeated exposure is necessary to achieve a level of comfort with information systems and believe it will be crucial for both faculty and staff to continue regular interactions with the system; therefore we are incorporating use of the system into every semester of our undergraduate curriculum before moving into the graduate program with UCARE AES.

To date, it has been implemented into the curriculum for 3 (of 4) semesters of the undergraduate nursing program, although only Phase I is the focus of this paper. First semester students use UCARE AES to learn how to document individual "patient" assessments. Second semester students use the system to locate data in a medical record, synthesize information from multiple sources, and make clinical decisions based on the synthesized data. Third semester students currently use the system for pediatrics and maternity nursing. In pediatrics students assess scenarios on patient simulators (manikins) and obtain the "patient's" history from UCARE AES. In maternity students use UCARE AES to review a case study, prepare care plans, enter nursing orders, and practice reading fetal monitoring strips. All third semester students use a form

called SBAR (situation, background, assessment, and recommendation) to practice communication with nurse and physician providers. The implementation team is presently meeting with fourth semester faculty and mapping their curriculum to UCARE AES. As stated, all semesters of the undergraduate program will be live by May 2007.

Acute care nurse practitioner faculty are anxious to incorporate UCARE into their curriculum. They will be added as soon as the baccalaureate curriculum implementation is complete. For the nurse practitioner students, curricular content will be matched to the "experienced nurse" competencies of Staggers, Gassert, and Curran. The School of Pharmacy has already expressed a desire to be added to UCARE AES. The discussions with pharmacy are pointing to the use of collaborative and interdisciplinary case studies that focus on pharmacotherapeutics.

Faculty are becoming more comfortable using the system as their experience increases. In general, faculty are beginning to recognize the value of teaching electronic information management concepts and are continually verbalizing how they think they can expand the use of UCARE in their curriculum. To help faculty adopt more of their learning activities to UCARE AES, we anticipate adding an informatics student in the role of a teaching assistant (TA) to the implementation team to work with clinical teaching staff to assist them in using the system. The TA will also be available to work with students who may be uncomfortable using the system. Midway through the first semester we added a Systems Analysis position to the team. That individual helps to interface with Cerner to resolve any problems that arise and to request needed changes to the system.

Conclusions

Since UCARE AES has been added to the undergraduate curriculum it is clear that students are gaining competence in using information technology and in doing electronic data management. This will help with the development of sustainable health information systems. In addition undergraduate students are becoming aware of nursing informatics as a field and some have expressed interest in specialization in this area at the graduate level. We look forward to being able to report data that is being collected throughout the project as the remaining phases of the implementation are completed.

References

[1] WCED: UN World Commission on Environment and Development. Our Common Future. Oxford: Oxford University Press, 1987. Available online at http://www.are.admin.ch/are/en/nachhaltig/international_uno/unterseite02330/

[2] Madani S and Aronsky D. Factors Affecting the Sustainability of Information Technology Applications in Health Care. AMIA Annual Symposium Proceedings, 2003: 922. Available online at http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1479905

[3] Yasnoff WA, Humphreys BL, Overhage JM, Detmer DE, Brennan PF, Morris RW, Middleton B, Bates DW, and Fanning JP. Consensus Action Agenda for Achieving the National Health Information Infrastructure. *J Am Med Inform Assoc*. 2004;11:332–338.

[4] Gassert CA. The Challenge of Meeting Patients' Needs with a National Nursing Informatics Agenda. *J Am Med Inform Assoc.* 1998; 5(3): 263-268.

[5] Kinder G, Cashman SB, Seifer SD, Inouye A and Hagopian A. Integrating Healthy Communities concepts into health professions training. *Public Health Rep*. 2000 Mar–Jun; 115(2-3): 266–270.

[6] Staggers, N, Gassert, CA, and Curran, C. Informatics Competencies for Nurses at Four Levels of Practice. *J of Nurs Educ.* 2001; 40(7): 303-316.

[7] Staggers, N, Gassert, CA, and Curran, C. A Delphi Study to Determine Informatics Competencies for Nurses at Four Levels of Practice. *Nurs Research.* 2002; 51(6): 383-390.

[8] Ornes, LL and Gassert, CA. Computer Competencies in a BSN Program. *J of Nurs Educ.* 2007; 46(2): 75-78.

[9] Gugerty, B. Nursing at a Crossroads – Education, research, training, and informatics. *J of Healthcare Information Management,* 2007, 21(1): 12-14.

[10] McDowell, DE and Xiping, MA. Computer literacy in baccalaureate nursing students during the last 8 years. *CIN*, 2007; 25(1): 30-36.

[11] Ivanitskaya L, O'Boyle I, and Casey AM. Health Information Literacy and Competencies of Information Age Students: Results From the Interactive Online Research Readiness Self-Assessment (RRSA). Med Internet Res. 2006 Apr–Jun; 8(2): e6. Published online 2006 April 21. doi: 10.2196/jmir.8.2.e6 at http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&rendertype=abstract&artid=1550696

[12] Staggers N, Gassert CA, and Skiba DJ. Health Professionals' Views of Informatics Education: Findings from the AMIA 1999 Spring Conference.J Am Med Inform Assoc. 2000; 7(6): 550-558.

[13] TIGER. The TIGER Initiative: Evidence and informatics transforming nursing: 3-year action steps toward a 10-year vision. Available online March 20, 2007 at http://www.tigersummit.com.

[14] Warren, JJ, Connors, HR, and Weaver, C. Clinical Information Systems and Critical Thinking: Outcomes of an Innovative Education Strategy. NI2003 Proceedings. 2003.

[15] Warren, JJ, Connors, HR, Weaver, C. and Simpson, R. Teaching Undergraduate Nursing Students Critical Thinking: An Innovative Informatics Strategy. NI 2006 Proceedings. 2006: 261-265.

Address for correspondence

Carole Gassert can be contacted at carole.gassert@nurs.utah.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

ROC van Twente: Nursing Education in Care and Technology

William TF Goossen[a,c] Anneke TM Goossen-Baremans[b], Laura Hofte[c], Bert de Krey[c]

[a] *Results 4 Care, Amersfoort, the Netherlands*
[b] *Acquest Research, Development and Consulting, Koudekerk aan den Rijn, The Netherlands*
[c] *ROC van Twente, Enschede*

Abstract

The ROC van Twente offers nursing education at the diploma level (MBO), and is innovating the program to include a major /minor structure for education about care and technology. In order to achieve this, a new position was created: the Master Docent, Care and Technology. The task of the master docent includes development of education for nursing about technology, multidisciplinary cooperation, and service to health care institutions among others.

The first development concerns a module about electronic patient records, standards, and semantic interoperability for continuity of care. The module is delivered to nursing students and to students from the information technology department, who work jointly in 'development teams'.

This paper describes the background, the development of the educational material and program, and the core content of the module. The core content are the care information models that link clinical materials with health care information standards. The program has started end November 2006. At the Medinfo 2007 conference the results of the course for the first group of about 40 students will be presented.

Keywords:

health informatics education, HL7, vocabulary, standards, continuity of care, electronic patient records

Introduction

National infrastructures for health care information exchange present a challenge to clinicians to adopt many standards. However, clinicians usually did not get education about standards, and even in this age, the curricula of schools of health professionals usually do not include such topics in the overcrowded programs, despite the growing need for it.

Ageing populations and increasing numbers of people with chronic diseases change the focus of health care in such a way that the application of technology becomes essential. Example technologies applied include home care technology, video surveillance, smart homes, and information and communication technologies such as telehealth and electronic patient records (EPR). Increasingly, these technologies are integrated with each other, for instance based on internet standards. Thus, new kinds of healthcare emerge and for health care professionals there is an imperative to make sensible decisions about what technology to introduce in the care environment. The integration of technologies and the changes in health care delivery depend more and more on standardisation and quality assurance.

This paper discusses how this approach to development of electronic messages and EPR is chosen as the topic for nursing education and for the education of application developers. A new educational program is introduced in this area, currently offered by the school of nursing and the school of technology of the ROC van Twente (Regional Educational Centre of Twente).

Background

In the Netherlands, the activities of the National ICT institute for Health Care (www.nictiz.nl) [1] lead to the emergence of standards for electronic message exchange and development of electronic patient records [2]. The use of the EPR will be required by law in the near future. Thus there is a need for all health professions to learn using it, and to be able to support the development of EPR that address the clinician's needs.

The Netherlands have based their 'information for health care' strategy on the message paradigm, applying the international Health Level 7 version 3 (HL7 v3) standard for the safe exchange of patient information to authorized users via a national infrastructure. HL7 v3 has been used in about 20 projects now as a method to determine (clinical) user needs, modelling these needs, and implement clinical content in electronic messages. In addition, several vendors successfully base their electronic patient record systems on the HL7 v3 models.

A key part of the developments include the HL7 v3 messages for continuity of care: this is the care provision domain [3]. This care provision message is meant for referral, acceptance, record exchange, discharge summaries and so on. It is a generic structure covering a standard way to identify sender and receiver, the patient, the purpose of the message and the expression of supporting clinical data. Clinical details are expressed in the 'care statements' or 'clinical statements'. The care statements themselves can vary significantly, but the way they are

included in the message is consistent over different clinical domains. This has lead to the development of care information models [4, 5] that standardize clinical content in such a way it can be used and re-used in the HL7 v3 care provision messages [5]. The purpose of this approach is to realise semantic interoperability between health care information systems and technologies. Semantic interoperability is considered such that professionals receiving patient data electronically, clearly understand the meaning of the message and can adequately continue the required care.

The school of nursing in Twente is timely with this approach due to the fact that the Twente region is a national pilot for the electronic patient record [7]. Health care providers in this region have requested more education in the area of care and technology, specifically about the use of information and communication technology. The program described below has been developed with input from representatives of the health care providers. The program is considered a try out for both education within the school of nursing, and for continuing education of the existing nursing workforce. These students from the school of nursing of ROC van Twente will eventually work as nurses, but with additional knowledge and experiences with the EPR.

Design of the educational program

The master docent, care and technology

The ROC van Twente is positioned in the east of the Netherlands. Twente is the front runner for the national implementation of the Nictiz spearhead projects: the medication record and the general practitioner to general practitioner record exchange. Therefore there is a perfect situational context for education. Students will be confronted with the developments in their practical / clinical traineeships.

However, for a school of nursing it is difficult to gain immediate expertise to start participation. Therefore a new teacher role was established: the first master docent for diploma level education. The master docent has responsibilities to innovate the education, in this example about care and technology, to bring in knowledge and experience, in this example built up on many projects in health standards [7], to involve the teachers, in this example via a project team, and to deliver service to health care facilities in the Twente region. Service to health institutions is delivered for instance via participation in requirements gathering workshops and traineeships.

To start the developments, a choice was made to keep it simple in the beginning, but at the same time take an example technology that is innovative and that relates to the national and regional developments of information and communication technology in health care. Thus, semantic interoperability was chosen as the leading principle to start the developments. Health care agencies in the region agree that they require nurses with skills and knowledge to participate in development of patient record systems and messages.

The care and technology module

The student groups participating in the minor care and technology include about 25 students from the school of nursing and about 15 students of the application developers program of the school of technology of the ROC van Twente. The teacher team is a multidisciplinary team of nurse educators and information technology educators.

An overall goal for the module that started November 23 is that both student groups, each from their own perspective, understand the process of determination of information requirements, standardization and development and implementation of electronic patient record systems and electronic messages. The care information models are intended as a framework that bridges the often existing communication gap between system users and system developers.

The minor program serves as a differentiation within the nursing and within the technical education. Therefore it is assumed that students do know the basics of nursing care and have experiences in traineeships before entering the minor program. For the technical students, an equivalent background in systems life cycle and methods applied is expected.

Specific learning objectives / required competencies include:

- Multidisciplinary cooperation
- Communication and active participation
- Analysis of the need for care (nursing students)
- Analysis of information needs (application developers).
- Development of a care information model that includes purpose, description of variables, codes, HL7 v3 model, and technical data specification.
- Development of functional requirements for a electronic patient record system for continuity of care.

In total, the program consists of 12 weeks of education. A total of 4 contact hours per week is presented in small working groups with a mix of nursing and technology students. In addition, the students need an average of 4 hours a week for reading and preparation of the teamwork. The program is presented in Table 1.

The first five weeks are about the need for electronic patient records and messages, and the content of the messages. Then the application developers fall back on their normal program, and use time in between to work on the functional requirements and system design. In the meantime the nurse students have a clinical traineeship. For six of the nursing students, the traineeship involves participation in a nursing system development. In this particular setting, the ROC van Twente, a home care agency, and a vendor work together to create a new traineeship.

After about 16 weeks, both the nursing students and the application developers come back to school and continue another 7 weeks of education on this module. These 7 weeks deal with the subjects presented in Table 1. They include continuity of care and requirements for electronic patient records and messages. Further, the development of a care information model is a core element. The students

finish the program with a presentation of requirements, design and examples. Teaching materials include a module and reading materials, based on the work for the national information and communication infrastructure.

Software: stroke care record system

Another teaching tool is the electronic stroke care record system by Portavita [8]. Portavita is a vendor that agreed to have their software for stroke systems, still under development, made available for the students of ROC van Twente. This is a cost neutral arrangement, where both parties benefit: ROC van Twente gets access to a professional clinical information system for education, the vendor gets exposure and feedback on the system, and educational materials developed around the system become available for clients.

Table 1 – Overview of the program for Care and Technology

Week 1. The need for the electronic patient record (EPR) and standards.	
Getting acquainted with each other. Nurses: describe what nursing needs in the EPR. Technicians: support the nurses and apply methods for requirements gathering and functional design. Determine how to cooperate for 12 weeks.	
Week 2. Analyse information in care	
Nurses: explain which data are required in a stroke care record system. Technician: interview the nurses to get the requirements for a system.	
Week 3. Care information model 1	
Study the care information model structure and start with making one example. Nurses the clinical and terminology part, technicians the model and technical specification.	
Week 4. Care information model 2	
Develop a care information model, including clinical, terminology, HL7 v3 model and technical specifications.	
Traineeship intermezzo	
Nurse student	**ICT student**
Study the existing paper based methods for continuity of care during traineeship	Develop a functional design for a nursing record system
Week 5. Continuity of care for stroke	
Study needs for continuity of care and describe processes, professionals involved, roles, tasks and activities, **information** to be exchanged and apply UML modelling.	
Week 6. Continuity of care record	
Prepare a continuity of care record for stroke patients: contents, standards, sequence diagrams and functional design.	
Week 7. Review existing materials	
Select all relevant care information models from the repository www.zorginformatiemodel.nl	
Week 8. Classifications and codes	
Apply coding from standardised nursing and health vocabularies	
Week 9. Evaluate design against existing system	
Compare the functional design with existing system for stroke care	
Week 10. Functional Design	
Discuss the functional design Technology student presents to working group and makes final adjustments.	
Week 11. Preparation	
Prepare a presentation for the whole class and for teachers Criteria include: 1) Agenda and minutes, 2) Even distribution of the work in the small group, 3) The subject of the presentation has been negotiated with and approved by the teacher, 4) Apply presentation software, 5) Include the following: care information model, functional requirements, evaluation of group work, functional design	
Week 12. Final assignment	
Present as workgroup the results in public	
Evaluation	Evaluate the course and work

Care information models as a core topic

In order to have concrete materials available that are manageable for the students, the use of the care information models [5] is taken as the lead during the minor program. The care information models serve as a reusable building block within the framework of HL7 v3 Care Provision messages [2, 3, 5].

Care information models combine different standards materials and create valuable content for intelligent semantic interoperability [5]. They function as a communication bridge between clinicians and technicians and

facilitate inputs into the technical development of electronic messages and EPR systems.

The document structure for the care information models consists of meta-information, detailed description of the clinical instrument, and the reason for its application in practice [5]. It specifies clinical care using professional evidence, uses standardized terminology and coding, uses standard (HL7 v3) information models, and specify at the detailed level the technical requirements for the clinical content. Thus, the technical implementation according the HL7 v3 message and data specification are included via mapping tables, which are useful for EPR development as well. In most documents, one for every item of clinical activities, observation, or instruments, the current components include, in a recently revised format, the following components [5, 9]:

1. Version management and authorship
2. Explanatory introduction about the use of care information models
3. Aim of the instrument, index, scale, act, or observation
4. Scientific foundations / evidence base or other foundation such as guidelines
5. Description of variables / data items / values
6. Working instructions for practice
7. Interpretation guidelines for the results
8. Information on the topic relevant for care process
9. References / acknowledgements
10. An example of the instrument (when available)
11. HL7 v3 message model and description
12. Mapping table from domain to standardized terminology and to HL7 v3 domain message model
13. XML message example (extensible markup language)
14. Copyright issues, such as licensing of source materials, allowed use of care information models
15. Screen designs / screen shots for the instrument
16. Remarks, e.g. if a Dutch version is different from English version of an instrument
17. Contact information: how to contact the author(s)
18. Disclaimer

A current overview of the 90 care information models, in the earlier – less complete – format, is available at the website: www.zorginformatiemodel.nl. [4]

Students work together in small groups to complete one (draft version) of a clinical relevant topic and specify it according to the above format for the care information model and complete the HL7 v3 specification.

Exam

The exam of the module consists of a presentation by the students of the design of a system for continuity of care, based on the professional and information standards. Nursing students must underpin their requirements from a patient care perspective. Application development students must substantiate their part with the analysis, modelling and design of a system (component) that meets nurses' requirements.

Future plans

Currently the module is taught to the two student groups. However, once the module is delivered, more work for the master docent and the project group is waiting. The following new developments are on the agenda for 2007 and 2008.

1. ROC van Twente wants to integrate the module to other health care providers' educational programs.
2. The content and assignments will be put into an electronic learning environment, thus making it available to students wherever they are.
3. The minor program will be made available for continuing education for the health care agencies in the region Twente.
4. Other ROC's (diploma schools) have joined or will join and a wider spread of the materials and the education on these subjects is in preparation.

Discussion and conclusion

National and regional developments of information and communication technology in healthcare, such as electronic patient records and electronic messages, are emerging in order to deal with the changes in the health situation of the Dutch population. Due to results of standardisation efforts, useful materials become available for education. The ROC van Twente decided to develop a new role, the master – docent care and technology – to assist in establishing an innovative program for nursing students and technology students.

The innovation includes several challenging areas. The first is a new kind of content: the sensible use of technology in the care environment. Secondly, the application of clinical, vocabulary, message and technical standards for exchange of information for continuity of care, based on a well established format of care information models. Thirdly, the use of electronic record software for stroke patients currently under development. Fourth includes the multidisciplinary student groups: nurses and technical developers, handling the same problem each from their own perspective. Fifth, a minor program within the existing education, indeed with a traineeship for currently few nurse students with a focus on system development. Finally, a project group from teachers, representatives of health agencies, and experts working together to achieve this.

It is an exciting area of developments and we will be proud to present the results of the first course during the conference.

Acknowledgments

We want to thank the participation of all involved teachers and heads of departments of ROC van Twente, and the members of the project team from the health institutions in the Twente

Region. Nelleke Plaisir of Acquest for her contributions to the module and reading materials.

References

[1] NICTIZ Nationaal ICT instituut in de Zorg. NICTIZ website. Available at www.nictiz.nl. Accessed November 2006.

[2] Goossen WT, Jonker MJ, Heitmann KU, et al. Electronic Patient Records: Domain Message Information Model Perinatology. *Int J Med Inf.* 2003; 70 (2-3):265-276.

[3] Health Level 7. Message standards. WWW documents. http://www.hl7.org/ Accessed November 2006.

[4] NICTIZ, Acquest and Portavita. Care Information Models. www.zorginformatiemodel.nl. [Webdocuments, accessed November 2006].

[5] Goossen WTF. Intelligent semantic interoperability: integrating knowledge, terminology and information models to support stroke care. In: Park HA, Murray P, Delaney C (Eds). Consumer-Centered Computer-Supported Care for Healthy People. Proceedings of NI2006. Amsterdam etc. IOS Press, 2006: pp. 435-439. Studies in Health Technology and Informatics, volume 122.

[6] Information Technology in de Zorg in Twente (2007). www.izit.nl [Web documents accesed March 2007]

[7] Goossen W. Representing clinical information in EHR and message standards: analyzing, modelling, implementing. HINZ Primary Care and Beyond: Building the e-Bridge to Integrated Care. Eds: Christopher Peck & Jim Warren. Publisher: Health Informatics New Zealand (HINZ) & Health Informatics Society of Australia Ltd (HISA), 2006.

[8] Portavita (2006). www.portavita.nl [Webdocuments. Accessed November 2006].

[9] Van der Kooij J, Goossen WTF, Goossen-Baremans ATM, Plaisier N. Evaluation of Documents that Integrate Knowledge, Terminology and Information Models. In: Park HA, Murray P, Delaney C, (Eds). Consumer-Centered Computer-Supported Care for Healthy People. Proceedings of NI2006. Amsterdam etc. IOS Press, 2006: 519-522.

Address for correspondence

William Goossen, Results 4 Care, De Stinse 15 3823 VM Amersfoort, the Netherlands, williamtfgoossen@cs.com, results4care@cs.com

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Multiple Measures of Provider Participation in Internet Delivered Interventions

Thomas K. Houston[a,c], Ellen Funkhouser[b,d], Jeroan J. Allison[c,d], Deborah A. Levine[a,c], O. Dale Williams[c], Catarina I. Kiefe[c,d]

[a] Division of General Internal Medicine, University of Alabama at Birmingham
[b] School of Public Health, University of Alabama at Birmingham
[c] Division of Preventive Medicine, School of Medicine, [d] Birmingham VA Medical Center

Abstract

Evaluation of Internet-delivered continuing education for health care providers requires appropriate consideration of their level of participation. To fully assess participation requires multidimensional measures, including factors such as the volume of participation (page views), frequency (visits), variety (components accessed by each provider), and duration (months of activity). We calculated crude and refined (adjusted for study design) measures and then compared these measures across three longitudinal Internet-delivered continuing education to health care providers (N = 429). We found that participation varied across study, varied by factor and varied by specific measure. Correlation between crude and refined measures within a factor and across factors differed significantly. Participation assessment of internet-delivered interventions varies by the selection of measure and factor. Further research assessing the potential for these measures to predict intervention effectiveness is needed.

Keywords:

internet, continuing education, educational measurement, quality of health care

Introduction

The Internet has promise as a new tool to increase translation of research into clinical practice.[1] If proven effective, low-intensity Internet-delivered continuing education interventions have a high potential for dissemination. These interventions seek to improve provider knowledge, motivation, and behavior and are often multi-modal with access to case-based education, decision support, and patient education materials. Many such interventions have the ultimate goal of improving quality and enhancing safety. Although evaluations of Internet-delivered continuing education for healthcare providers are increasing[2-8], no standard method of evaluating these interventions exists. To more fully understand the impact of these interventions, investigators need to measure participation, including tracking of all provider "encounters" with the program.

Tracking provider-intervention encounters, however, is challenging. Web tracking software most frequently tracks page views and visits.[9, 10] However, all page views may not be of equal value. A frequent goal of website developers is to create "stickiness" - a measure of repeat usage of the intervention.[11] Thus, investigators might want to assess frequency of use per month enrolled in a longitudinal intervention. A combination of measures may best represent actual participation.

Few previous studies of Internet-delivered continuing education have measured differences in participation. Also, it is unknown how different measures of participation are correlated. As a first step toward developing a scientific approach to measuring participation, we propose a four-dimensional model. The general factors include volume, frequency, variety of components accessed, and duration of access. Within this model, we developed crude and refined measures of each factor and then compared these measures using data from three longitudinal Internet-delivered continuing education interventions. Our goal was to determine how the measures varied by study and to assess the correlation of crude and refined measures within factor and across factor.

Methods

Study design

Our group is conducting three separate group-randomized trials of Internet-delivered educational interventions to improve practice patterns of health care providers. Funded by the U.S. National Institutes of Health (NIH) and Veterans Affairs Health Services Research & Development (VA), these trials have similar designs and objectives but different target populations of providers and disease foci. The National Heart, Lung and Blood Institute (NIH) MI-Plus, conducted in two southern states, and VA MI-Plus, a national study, are parallel interventions targeting different provider populations: Medicare primary care providers in two Southern states and clinic-based primary care VA providers, respectively. Both MI-Plus studies seek to improve care for ambulatory post-myocardial infarction patients with multiple co-morbidities. Funded by the National Institute on Drug Abuse (NIH) DentalTobaccoControl.net (DTC) targets dental providers and seeks to improve tobacco cessation practice patterns in dentistry.

We prospectively tracked participation by 429 healthcare providers enrolled in the intervention arms of these trials. Users (private practice physicians, nurse practitioners, physician assistants, dentists, and hygienists) were recruited to the websites through mailings, phone calls, and emails and enrolled over multiple months ranging

from six to twelve, depending on the study. For this analysis, tracking data for each study was collected through a specific calendar month in 2006 (March for NHLBI MI-Plus, August for VA MI-Plus, and September for DTC). As enrollment was rolling, total months of enrollment varied for individual providers (mean months of enrollment was 14 (SD 5) for NHLBI MI-Plus, 11 (SD 4) for VA MI-Plus, and 12 (SD 4) for DTC). Once enrolled, all providers received scheduled email reminders, including notifications of new content, to encourage participation. All providers received continuing education credits specific to their specialty (medical or dental) and a certificate of appreciation for their participation. Additional incentives were provided in the NHLBI MI-Plus study (access to online journals and a textbook) and VA MI-Plus (subscription to the *Medical Letter*). Each study was approved by the appropriate Institutional Review Board.

Intervention descriptions

The core of all three interventions was case-based educational programs using interactive, web-based modules with tailored feedback based on responses to questions. During the tracking periods, the content and number of cases varied by study (6 for MI-Plus, 8 for VA MI-Plus, and 3 for DTC). All three studies had an accompanying "toolbox" with practice tools and patient educational materials that could be downloaded. The two MI-Plus studies included 1) a literature watch segment updated at intervals with reviews of the literature and 2) a guidelines component with summaries of current guidelines applicable to post-myocardial infarction patients. The literature watch and guidelines of the MI-Plus studies were analogous to the headlines and library components of the DTC study. Feedback of performance data with peer comparisons was provided to NHLBI MI-Plus providers. In the DTC study, testimonials of provider's success in encouraging smoking cessation were included. For this analysis, we focused on intervention components that were consistent across the three studies.

User authentication was required for all providers as they logged onto the interventions. Thus, we used server tracking logs linked to visit to calculate the measures of participation. The log included an individual user identification number and was tagged with date and time.

Measures of participation

As noted above, we propose a four-dimensional model to evaluate participation. The four factors are 1) volume, 2) frequency, 3) variety of components accessed, and 4) the duration of activity. For each factor, we developed crude and refined measures for each factor. Volume measures included total number of page views (crude) and the refined measures (number of page views per visit and number of pages per month). Frequency measures included total number of visits (crude) and number of visits per month (refined). Variety measures included number of components accessed (crude). Because the central component of the three interventions was the cases, we also created variety measures specific for the cases: number of case modules completed (crude), and mean percent of case modules completed (refined). Duration measures included 1) number of months actively participating (crude), defined as months from first to last logon, and 2) mean percent of enrolled months (refined), defined as the proportion of enrolled time known to be active (months active/months enrolled) because of the variation in potential enrollment across studies. We defined five categories of providers (private practice physicians, VA physicians, VA nurse practitioners/physician assistants, dentists and hygienists).

Other measures of participation could be calculated. Specifically, our measures focused mostly on counts, not session time, or time per webpage. Using session time is challenging as our providers rarely logged off, but would simply close the site so that session time continued indefinitely. Also, frequent outliers for session time existed (over five hours of activity in a single visit) suggesting that providers would just leave the page open and go to another task. Thus, we have not used session time in this analysis.

Statistical analysis

We calculated the means and standard deviations of each measure. We assessed differences in participation measures by type of provider using t-tests. Because of the multiple comparisons, we chose a significance level of = 0.01. Then, we evaluated participation in each intervention component by provider group and assessed differences between provider groups using chi-square tests. Finally, using participants in all three studies, we assessed pair-wise correlations between measures within and across factors using Spearman's rank correlation coefficient. We then repeated this analysis for each individual study again with = 0.01.

Results

We recruited 429 providers from 344 practices. These included 108 private practice primary care physicians in the NHLBI MI-Plus study, 193 VA primary care providers (125 physicians and 68 physician assistants or nurse practitioners) in the VA MI-Plus study, and 128 private practice dental providers (68 dentists and 60 hygienists) in the DTC study. Across studies, the mean number of months since enrollment was 12.2 months (SD 4.5).

Results by study

Overall, across measures of volume, frequency, variety, and duration of participation, values tended to be higher for the VA providers, both physician and non-physician, compared with private practice physicians and dental providers (Table 1). The mean visits per month among VA physicians, 0.62 (SD 0.52), was twice that of private practice physicians, 0.29 (SD 0.22), P = 0.001; the mean visits per month for VA Nurse-practitioners/Physician Assistants was even higher, 0.85 (SD 1.5), compared with private practice physicians, P = 0.0002. The point estimate for page views per month was highest among dentists in the DTC study (19.5) but the standard deviation was quite wide (106.6), thus this estimate was not significantly different than the estimates amongst the other providers.

Variety of access as measured by number of components was lowest among the dental providers [1.5 (SD 1.2) for dentists and 0.97 (SD 0.89) for hygienists] compared with the other studies, but as measured by proportion of cases completed, participation in the DTC study was similar to the VA MI-Plus study. Duration of participation ranged from a mean of 7.9 months (SD 4.1) in NHLBI MI-Plus (the longest running study) to 2.8 months (SD 4.2) for hygienists.

*Table 1 - Mean Measurements (Standard Deviations) of Participation for Providers in Three Internet-delivered Intervention Studies**

	NHLBI MI-Plus	VA MI-Plus		DentalTobaccoControl.Net	
	Private Practice Physicians (n = 108)	VA Physicians (n = 125)	VA NP/PA‡ (n = 68)	Dentists (n = 68)	Hygienists (n = 60)
Participation Measures*	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
A. Volume					
Total number of Page Views	60.4 (51.4) [a,b,c,d]	100.3 (74.5) [a,e,f]	89.2 (73.3) [b,h]	35.1 (33.0) [c,e]	24.5 (25.0) [d,,h]
Number of Page Views Per Visit	16.4 (8.8) [c,d]	15.6 (8.2) [e, f]	15.6 (9.6)	9.7 (7.11)[c, e]	9.5 (6.0) [d,f]
Number of Pages Per Month	4.5 (3.8) [a,b,d]	9.3 (8.1) [a]	13.6 (30.4) [b]	19.5 (106.6)	2.6 (3.8) [d]
B. Frequency					
Total number of Visits	3.8 (3.2) [a,b,d]	6.7 (5.5) [a,e,f]	5.9 (5.4)	3.0 (3.7) [e,i]	2.5 (2.4) [d,f,i]
Number of Visits Per Month	0.29 (0.22) [a,b]	0.62 (0.52) [a]	0.85 (1.5) [b]	1.00 (4.0)	0.25 (0.27)
C. Variety					
Number of Components	2.3 (1.3) [a,b,c,d]	2.8 (1.2) [a,e,f]	2.7 (1.3) [b,g,h]	1.5 (1.2) [c,e,h]	0.97(0.89) [d,f,g]
Number of Cases	1.7 (1.6)	3.5 (2.5)	3.2 (2.7)	1.5 (1.3)	1.3 (1.2)
Mean Percent of Cases completed	28%(27) [a,c,d]	44% (32%)[a]	40% (33%)	50%(44%)[c]	43%(41%) [d]
D. Duration					
Months from First to Last logon	7.9 (4.1) [c,d]	7.0 (5.4) [e,f]	6.2 (5.5)	4.8 (4.9)[c,e]	2.8 (4.2) [d,f]
Mean Percent of enrolled months	51%(38) [d]	56% (37) [e,f]	53 (40)%	39% (38) [e]	24% (35) [d,f]

* Participation measured in four dimensions – volume, frequency, variety, and duration of access. Cells in the same row with the same superscript are statistically significantly different, p<0.01
Crude measures, ‡ nurse practitioners and physician assistants
§ Percent of enrolled months = (Months for first to last logon/total months since enrollment)

When comparing use of individual intervention components, the VA providers had high participation, private practice physicians had moderate participation and dental providers had lower participation (Table 2). The VA physicians participated in a higher percentage of the cases (84%), compared with the private practice physicians (71%; P< 0.01) and the dentists (68%;P<0.01). Notably, a much larger proportion of VA physicians accessed literature watch than private practice physicians (82% vs 38%, P=0.01). Similar results were seen comparing the VA non-physicians with all other categories. The dental providers had the lowest rates of accessing each of the four factors.

Correlation of measures

The correlation between individual measures varied from strong (0.93) to weak (0.11) (Table 3). In general, "pages per visit" was the least correlated with the other measures. Also, the duration measures had a higher number of moderate or weak correlations with the volume, frequency, and variety measures. Within each of the four factors, correlations also varied. Some of the weakest within-factor correlations were again seen for measures of volume. The pattern of strength of the correlations seen in Table 3 was not qualitatively different when the analysis was repeated for each individual study (data not shown).

Discussion

To document the effectiveness of Internet-delivered educational interventions that seek to translate research into practice, we need more randomized trials. To fully evaluate these trials, investigators will need to describe the level of participation in the intervention as a surrogate for exposure. We evaluated a number of easily calculated measures of provider participation in Internet-delivered continuing education interventions for healthcare providers and noted three general patterns of variation – overall patterns by study, patterns across study by similar measures, and patterns of correlation among measures.

Table 2 - Participation in Individual Components by Provider Type in Three Internet-based Interventions

	Private Practice Physicians	VA Physicians	VA NP/PA*	Dentists	Hygienists
	N = 108	N = 125	N = 68	N = 68	N = 60
Any Interactive Case	71%	84% [e,f]	81%	68% [e]	63% [f]
Literature Watch or Headlines	38% [a,b,d]	82% [a,e,f]	75% [g,h]	32% [e,g,i]	6% [d,f,h,i]
Guidelines/Library	52% [c,d]	44% [e,f]	50% [h]	12% [c,e]	3% [f,h]
Toolbox	63% [d]	69% [e,f]	63% [g]	45% [e,i]	23% [d,f,i]

Cells in the same row with the same superscript are statistically significantly different, $p<0.01$
* nurse practitioners and physician assistants

Overall, the measures suggested varying rates of participation by study, with VA providers having the highest participation. This may be because VA providers have relatively greater computer access. It may also be related to "horizontal" use of technology. All VA providers in the community-based outpatient clinics use computers to access an electronic health record; thus, use of the computer for Internet-delivered continuing education during routine workflow may be more feasible and acceptable than in the private practice physician and dentist offices where use of the computer for other purposes likely has considerably wider variation.

We also found that the crude measures had more statistically significant differences across studies compared with the refined measures. Our refined measures included adjustment for variations at the study level to account for differences in exposure to the intervention. For example, number of months of enrollment varied across the studies. Thus, the total time available for participation varied considerably. Also, total number of pages was greater in the MI-Plus studies compared with the DTC study. Thus, the point estimate for total pages favored the physician groups but "pages per month" for dentists was higher although not statistically significantly so. Conceptually, we feel the refined measures are likely more appropriate when comparing participation across studies because they adjust for differences in the interventions and may provide a more accurate comparison across studies.

Third, we found that the correlation statistics varied comparing the various measures both within dimensions and across the dimensions of volume, frequency, variety, and duration of participation. "Pages per visit" had the weakest correlations with the other measures. Duration of participation showed only moderate correlation with the adjusted measure of visits per month and pages per visit.

These measures demonstrate both the successes of the three interventions and some difficulties. All three interventions successfully maintained provider participation over a number of months. However, participation did wane, with most providers only participating for around 50% (or less for the DTC study) of the total time available. In the DTC study, providers less frequently accessed the non-case intervention components. We have used this information to increase marketing emails encouraging current DTC study participants to access the toolbox and other intervention content. Our efforts to maintain participant

*Table 3 - Spearman's Correlation Coefficients for Measures of Participation**

	a	b	c	d	e	f	g	h	i
Volume									
a. Total Pages	1.0								
b. Pages Per Visit	0.58	1.0							
c. Pages Per Month	0.89	0.57	1.0						
Frequency									
d. Total Visits	0.85	0.13	0.74	1.0					
e. Visits per Month	0.72	0.11	0.84	0.82	1.0				
Variety									
f. Number of Components	0.80	0.54	0.73	0.67	0.56	1.0			
g. Number of Cases	0.92	0.55	0.82	0.78	0.65	0.71	1.0		
h. Percent of Cases	0.78	0.45	0.68	0.68	0.577	0.55	0.90		
Duration									
i. Months from First to Last logon	0.72	0.17	0.79	0.54	0.53	0.53	0.63	0.56	1.0
j. Percent of enrolled months	0.72	0.19	0.62	0.79	0.65	0.51	0.64	0.58	0.93

* Spearman's rank correlation coefficient, bolded cells represent those Not significant ($p > 0.01$)
Percent of enrolled months = (Months for first to last logon/total months since enrollment)

intervention activity had varying effects, with wide variations in participation of individual providers within studies. For some measures, the standard deviations were equal or greater than the point estimates. Future studies may require targeted marketing strategies to individual providers or smaller targeted groups to enhance participation. Also, working with individual practices to understand barriers to participation and to facilitate access may be required.

Each of our four factors has some face validity, the measures demonstrate within study variation, and we have demonstrated that the measures differ across studies. This suggests that our measures capture differences in level of participation, as intended. Still, further analyses are needed to fully assess the multi-factorial model of participation that we propose. Additional analyses evaluating factors that may be associated with participation (number of computers, acceptability of technology, computer literacy) should be conducted to assess whether participation can be predicted. Most important, analyses need to be conducted to assess whether these measures of participation predict outcomes including changes in knowledge and performance, and which measures are better predictors of outcomes.

Prior educational literature suggests that approaches that provide education in small content amounts separated by time ("spaced approach") are more effective than large content approaches. Thus, one might hypothesize that frequency and duration of participation are both critical. However, these factors cannot be considered in isolation, as one might have high frequency participation in the intervention with low volume and low variety. In prior research, a model for web surfing – the foraging model – suggests that there are discreet patterns of participation[12]. Thus, the four factors might be combined to categorize users into certain patterns of participation. We have not attempted to combine these measures into a single participation factor because we feel that further empiric validation of criterion validity is needed first.

A recent manuscript measuring participation in a patient website[13] noted that participation is but one measure of engagement and that participation should be considered in the context of satisfaction with content, assessments of knowledge, and process measures that may potentially lead to the desired endpoints. We feel this also holds for provider interventions. To fully predict endpoints, measures other than participation will also be needed.

Our study was limited in that data from only three interventions were available. As noted above, this is a first evaluation of these measures, and includes only correlational evidence, but additional assessment of construct and criterion validity are the object of ongoing research. Differences in participation across studies may be due to variations in interest, time availability, computer experience, and ease and type of computer access, none of which we could assess in this analysis.

Conclusion

In conclusion, participation of providers in Internet-delivered interventions varies widely across multiple dimensions of measurement. Primary and refined measures of volume, frequency, variety, and duration discriminated across studies. Within study variation of these measures was also high, as noted by the high standard deviations. In our studies, these measures have been used to track ongoing participation, suggest modifications to website design and marketing, and will be useful in understanding the outcomes of these interventions. Our preliminary analyses add to the prior literature by identifying categories of participation measures, and begin to approach the complexity of assessing participation. Much further research is needed to develop and validate the science of measuring participation for Internet-delivered continuing education.

Acknowledgement

This project was funded in part by the National Institute on Drug Abuse (R01 DA017971), the National Institute of Dental and Craniofacial Research (U01 DE16747) and the National Heart, Lung, and Blood Institute, (R01 HL70786).

References

[1] Weingarten S. Translating practice guidelines into patient care : guidelines at the bedside. *Chest.* Aug 2000; 118(2 Suppl):4S-7S.

[2] Allison JJ, Kiefe CI, Wall T, et al. Multicomponent Internet continuing medical education to promote chlamydia screening. *Am J Prev Med.* Apr 2005;28(3):285-290.

[3] Curran VR, Fleet L. A review of evaluation outcomes of web-based continuing medical education. *Med Educ.* Jun 2005;39(6):561-567.

[4] Wutoh R, Boren SA, Balas EA. eLearning: a review of Internet-based continuing medical education. *J Contin Educ Health Prof.* Winter 2004;24(1):20-30.

[5] Stewart M, Marshall JN, Ostbye T, et al. Effectiveness of case-based on-line learning of evidence-based practice guidelines. *Fam Med.* Feb 2005;37(2):131-138.

[6] Fordis M, King JE, Ballantyne CM, et al. Comparison of the instructional efficacy of Internet-based CME with live interactive CME workshops: a randomized controlled trial. *Jama.* Sep 7 2005;294(9):1043-1051.

[7] Butzlaff M, Vollmar HC, Floer B, Koneczny N, Isfort J, Lange S. Learning with computerized guidelines in general practice?: A randomized controlled trial. *Fam Pract.* Apr 2004;21(2):183-188.

[8] Chung S, Mandl KD, Shannon M, Fleisher GR. Efficacy of an educational Web site for educating physicians about bioterrorism. *Acad Emerg Med.* Feb 2004;11(2):143-148.

[9] Creese G, Burby J. *Web Analytics Key Metrics and KPIs (Version 1.0)*. Washington, DC: Web Analytics Association; 2005.

[10] Peterson E. *Web Analytics Demystified: A Marketer's Guide to Understanding How Your Web Site Affects Your Business*. Seattle: Celilo Group Media; 2004.

[11] Khalifa M, Limayem M, Liu V. Online consumer stickiness: a longitudinal study. *Journal of Global Information Management.* 2002;10(3):1-14.

[12] Pirolli P, Card S. Information foraging. *Psychological Review.* 1999;106(643-675).

[13] Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in Web-based health behavior change programs. *J Med Internet Res.* 2006;8(3):e15.

Address correspondence and reprint requests to:

Thomas K. Houston, MD MPH
510 20th Street South, FOT 720B
University of Alabama at Birmingham
Birmingham, AL 35294-3407
thouston@uab.edu
205-934-7997, Fax: 205-934-7797

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Medical Students' Knowledge and Perceptions of e-Health: Results of a Study in Sri Lanka

Sisira Edirippulige[a], Rohana B. Marasinghe[b], Anthony C. Smith[a], Yoshikazu Fujisawa[c], Walisundara B. Herath[b], M. T. M. Jiffry[b], Richard Wootton[a]

[a] Centre for Online Health, University of Queensland, Australia
[b] Faculty of Medical Sciences', University of Sri Jayewardenepura, Sri Lanka
[c] Niigata University of Health and Welfare, Japan

Abstract

The present study investigates the knowledge, perceptions and attitudes of medical students in Sri Lanka in regard to e-health. We also examined the barriers which impede them to develop knowledge and skills in e-health within their medical curriculum. A questionnaire focusing on the knowledge, attitudes and expectations of medical students towards e-health was distributed to all final year students (n=136) at the Faculty of Medical Sciences, Sri Jayewardenepura University, Sri Lanka. Response rate was 74%. 43% of respondents stated that they were familiar with the term e-health. 51% rated their knowledge of e-health applications as minimal. 88% admitted that they had no e-health education or training of any kind. Over 80% of all respondents thought that e-health had an important role to play in the current and future health sector, particularly in developing countries. Our survey revealed that respondents had very poor access to computers and Internet use was rare. 77% of respondents admitted that they were not provided with systematic knowledge and skills in e-health through their medical curriculum and identified the absence of formal education in e-health as a serious shortcoming.

Keywords:

medical education, e-health, curriculum development, developing countries

Introduction

The healthcare industry in general is under growing pressure to provide services more efficiently and economically [1]. As far as developing countries are concerned, problems and issues associated with the health services are even more daunting [2]. Extreme poverty has prevented governments in developing countries from funding health adequately resulting poor state of health in the populations. Poor health infrastructure, inadequate facilities and the shortage of healthcare professionals are characteristic features of the health sector in developing countries. Rural and remote communities – where the majority of the population live in developing countries - are particularly deprived of health services. Addressing health issues in developing countries has become a global concern [3]. However, there is no quick remedy to health problems in developing countries as the improvement of health infrastructure and the increase of the number of health professionals require long-term investments and policy implementations [4].

In this context, e-health has been identified as one alternative to provide improved services and specialist care in developing countries [5]. The potential of e-health applications, i.e. the delivery of health services across a distance by using information and communication technologies (ICT) is being recognised for its potential, especially for the delivery of health services to rural and remote areas [6]. International organizations such as the United Nations (UN) and the World Health Organization (WHO) have acknowledged e-health as a potential alternative to address acute health needs in developing countries [7-8].

Among many other factors, the acceptance of e-health by health professionals is extremely important if this new modality of healthcare delivery is to become an integral part of mainstream healthcare. Knowledge, acceptance and enthusiasm to utilise e-health as an alternative way of service delivery by healthcare professionals, particularly by doctors would help facilitate the integration of e-health. Needless to say, developing countries are one of the most likely beneficiaries of e-health. Thus, knowledge and skills in e-health and keenness to use this tool by local healthcare professionals would undoubtedly help address at least some of the existing problems.

There is a growing body of literature showing that e-health has a role to play in contemporary healthcare [9-10]. Also literature is evident that e-health is useful for developing countries [11]. However, there is not many studies done to investigate the level of knowledge among health professionals [12]. A few studies show the level of IT knowledge in medical students [13].

Sri Lanka is a developing country according to number of indicators. Like other developing countries, health sector in Sri Lanka has been challenged by various problems. Total expenditure on health as a percentage of GDP in 2002 was 3.7% which is significantly lower than that is in

industrial countries. Overall access to health facilities for people in Sri Lanka is low. For example the number of hospital beds per 1000 population constitutes 2.9. The number physicians per 10,000 population is 4.11 [14].

Rural population in Sri Lanka constitutes 77% of the total population [15]. The most complete hospital facilities and highest concentration of physicians are in the urban areas, while many rural and remote areas suffer severe scarcity of health services. Thus the territorial disparity in health services is a characteristic feature in the island (Table 1). Emergency transport of patients especially in the countryside is still at a rudimentary level.

Table 1 - Number of medical specialists in urban and rural hospitals in Sri Lanka[1]

Speciality	Urban	Rural	Total
Cardiologists	18	0	18
Neurologists	9	0	9
Psychiatrists	15	0	15
Pathologists	23	9	32
Dermatologists	10	1	11
Radiologists	29	6	35
Microbiologists	16	0	16
Occupational therapists	46	1	47

[1] Source: Sri Lanka Government Health Web Portal, 2005, (http://www.health.gov.lk/)

Aims

The objective of this present study was to evaluate the knowledge, attitudes and perceptions of medical students towards e-health since their preparedness is a key to the success in implementing e-health in developing countries.

Methods

We designed and distributed a survey to assess the knowledge and attitudes of medical students towards the broad subject of e-health. The survey was distributed to all final year medical students (136) studying at the Faculty of Medicine, Sri Jayewardenepura University (SJU), Sri Lanka. Questions were divided into the following sections: demographic details, knowledge in e-health, relevance to future practice, the use of computers and the Internet and access to e-health education.

Results

Demographics

A total of 100 (74%) students completed the survey. 54% of respondents were female. The majority of respondents (about 91%) were between the age of 26-30 years and the remainder were between 23-26 years of age.

Knowledge of e-health

Nearly half of all respondents (43%) admitted that they were familiar with the term e-health. However, 51% of respondents described their knowledge and skills related to e-health as minimal while 22% were unsure. 86% of respondents had had no exposure to e-health education and/or training. 71% of respondents said they had never read any literature on e-health.

Relevance of e-health

Questions were asked to examine the perceptions of the students about e-health. About 86% of all respondents admitted that e-health will have an important role to play in the current and future health sector. Only 2% disagreed with that statement while 11% were not sure. Again 86% of respondents agreed with the fact that e-health will be useful in their future practice. Only a very small number admitted that e-health will have no use in their future practice. 78% of respondents admitted that e-health applications will improve their services. Majority of respondents (77%) believed that e-health would have particular relevance to developing countries and 85% agreed that e-health should be encouraged.

Use of computers and the internet

Several questions were asked to establish the knowledge and skills of the participants in computing and the level of the Internet use. The results of the survey showed that the availability of computers and the Internet for students was low. They admitted that the access to computers and the Internet was limited both at home and at the university. Only a very small number (3%) of students had frequent access to computers and the Internet. The majority of students (65%) used the Internet very rarely. Nonetheless a large number of students admitted that they were comfortable using computers and the Internet. Also 67% admitted that they had formal computer education and training. The majority of students expressed the desire to have better and more frequent access to computers and the Internet.

Access to e-health education

41% of respondents admitted that they had received no satisfactory knowledge of e-health through their medical program while 36% were not sure. While 79% of respondents suggested that e-health should be included in the medical curriculum and 56% thought that e-health must be offered as an elective. About 85% of survey participants suggested that e-health course must include a practical component to provide hands-on skills. More than half of respondents (64%) expressed their willingness to study e-health at post-graduate level.

Participants of the survey also identified the lack of appropriate educational programs, financial constraints, lack of

Table 2 - Information and Communication technology distribution

Countries	Main telephone lines per 100 persons	Residential main lines per 100 households	Monthly subscription as % of income per capita	Personal computers per 100 persons	Internet users per 10,000 persons	Internet hosts per 10,000 persons
Low income	2.9	11.4	14.1	0.6	62.2	1.0
Lower middle income	13.6	35.8	2.9	2.4	264.9	4.3
Upper middle income	22.7	59.8	2.0	8.2	992.6	78.7
High income	59.7	108.8	0.7	37.3	3992.9	1484.2
World	17.1	54.9	5.7	7.7	820.8	232.6
Africa	2.6	9.9	12.7	1.0	84.9	3.4
Americas	35.1	80.6	3.1	26.6	2164.3	1332.9
Asia	10.7	41.8	5.5	2.2	433.9	28.7
Europe	40.5	80.0	1.1	17.9	1804.5	191.5
Oceania	40.0	98.3	3.7	39.9	2771.6	885.2

Source: International Telecommunication Union, World Telecom Indicators 2002

sufficient access to technology and traditional methods of medical education as major barrier to develop systematic knowledge and skills in e-health.

Discussion

There are no quick solutions to the complex problems in the health sector in developing countries. Among others, e-health has been identified as one possible solution to address some of these problems. Under right circumstances new technologies can improve the quality of care and efficiency of services. Governments as well as private sector around the world have become aware of the potential of new technology. But the enthusiasm of policy makers and investments in infrastructure only cannot enable e-health to enhance health services. The expansion of knowledge and skills in e-health at grass-roots level (among health professionals) and their acceptance of these techniques are imperative factors for e-health to become sustainable.

Our survey revealed that although the majority of students were familiar with the term e-health their knowledge and skill to practice this modality was extremely limited. Indeed, the limited access to computers and the Internet is a serious barrier in developing countries. Unlike in industrialised countries, computers are still a luxury in the developing world. Presumably this barrier has limited the advantages they may gain from new technologies. This limitation also represents a significant factor preventing them from acquiring necessary knowledge and skills in e-health. The so called 'digital divide' is still a formidable barrier to be overcome [16].

Despite the fact that students have limited access to computers and the Internet, the majority of them are computer literate. In fact, there is a growing interest in computers and the Internet in the developing world [12]. The students found that one of the main barriers for them to develop appropriate knowledge and skills in e-health was the absence of formalised e-health educational components in the medical curriculum. There is a need to provide knowledge about the fundamentals of e-health, basic concepts and various applications with particular emphasis on low-cost e-health modalities. Such education must also include a practical component to provide medical students with necessary hands-on skills. This preparation would enable students to choose relevant applications in their own practice suitable for their own circumstances.

Conclusions

Health systems in developing countries can be a potential beneficiary of e-health applications. Not only local governments, but also the international organisations such as the UN and the WHO have identified the potential of ICT to address the health needs of developing countries. Efforts have been made to promote e-health in developing countries by investing funds, initiating projects and introducing technology and improving infrastructure.

However, in this effort, e-health education has been the least attended area. The knowledge, acceptance and enthusiasm of local health professionals, particularly doctors are vital if e-health is to be a significant component of mainstream healthcare. Cultivation of a positive attitude towards e-health requires systematic education. Students must be provided with a formalised e-health education within their medial and health curriculum to establish knowledge in basic concepts, terminology, various e-health applications, successes and failures in current practice. Such education must also include a practical component to provide hands-on skills.

Undoubtedly the impact of digital divide is still a serious problem for developing countries. Concerted efforts must

be made to enhance the access to technology in these countries. The potential of low-cost e-health applications in developing countries is still untapped. In order to use low-cost e-health modalities to their full capacity one must have an appropriate knowledge in e-health and understanding of local needs. Formalised e-health education embedded into medical and health curriculum is needed to enhance the knowledge and skills of local health professionals.

Acknowledgments

We thank the University of Queensland (New Staff Start-Up Research Fund) for funding this research. We also thank the medical students of Sri Jayewardenepura University for their participation in the study.

References

[1] Lee K, Fustukian S and Buse K. Health policy in a globalising world, Cambridge, Cambridge University Press, 2005.

[2] Carr D. Improving the health of the world's poorest people, Health Bulletin 1, Washington, DC: Population Reference Bureau, 2004.

[3] Wagstaff A, Claeson M. Millennium Development Goals for Health, Nutrition and population: Rising to the challenges, World Bank Publications, 2004.

[4] Akin JS, Birdsall N, De Ferranti DM. Financing health services in developing countries, World Bank Publications, 1987.

[5] Latifi R. Establishing telemedicine in developing countries: from inception to implementation, Amsterdam, IOS Press, 2004.

[6] Lakovidis I. E-Health: Current situation and examples of implemented and beneficial e-health applications, Amsterdam, IOS Press, 2004.

[7] United Nations, World Public Sector Report: E-Government at crossroads, United Nations Publications, 2003.

[8] World Health Organization, e-Health tools and services, needs of the member states, Report of the WHO Global Observatory for e-Health, 2006.Harrison JP, Lee A. The role of e-health in the changing health care environment. Nurs Econ 2006 Nov-Dec;24(6)283-8

[9] Reed K. Telemedicine: Benefits to advanced nursing practice nursing and the communities they serve. J Am Acad Nurse Pract 2005; 17: 176-80.

[10] Kirigia JM, Seddoh A, Gatwiri D, Muthuri LH, Seddoh J. E-health: determinants, opportunities, challenges and the way forward for counties in the WHO African Region. BMC Public Health, 2005 Dec 20;5:137

[11] Edirippulige S, Smith AC, Young J, Wootton R, Knowledge, perceptions and expectations of nurses in e-health: results of a survey in a children's hospital, Journal of Telemedicine and Telecare, 2006, 12 (Suppl. 3):S3:35-38.

[12] Gschwendtner A, Netzer T, Mairinger B, Mairinger T. What do students think about telemedicine? J Telemed Telecare, 1999;3(3):169-71

[13] Samuel M, Coombes JC, Mirinda J, Melvin R, Young EJW, & Azarmina P. Assessing computer skills in Tanzanian medical students: an elective experience, BMJ Public Health, 2004, available from http://www.biomedcentral.com/1471-2458/37.

[14] World Bank, Sri Lanka Home, Data and Statistics, available from http://www.worldbank.lk/WBSITE/EXTERNAL/COUNTRIES/SOUTHASIAEXT/SRILANKAEXTN/0,,menuPK:287056~pagePK:141132~piPK:141109~theSitePK:233047,00.html.

[15] Norris P. Digital Divide, Cambridge, Cambridge University Press, 2001.

[16] Smith R. Closing the digital divide, BMJ, 2003;326: 238.

Address for correspondence

Dr Sisira Edirippulige
Centre for Online Health
University of Queensland
Level 3 Foundation Building
Royal Children's Hospital
Herston Road HERSTON Q 4029
(Fax: +61 7 3346 4705)
Email: s.edirippulige@uq.edu.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Importance of Public Health Informatics: A Survey of Public Health Schools and Graduate Programs in the United States

Janise Richards, MS, MPH, PhD

*Centers for Disease Control and Prevention, Atlanta Georgia, USA**

Abstract

This paper examines the importance of data, information, and informatics to public health practice. Forty public health academicians from 40 schools and graduate programs of public health were interviewed. All agreed that informatics was important to public health practice. A qualitative analysis of their comments revealed their beliefs on the importance of informatics skills and knowledge to the practice of public health. The resulting comment groups varied from "some skills are more important than others" to "need all the skills." Eight "importance" comment groups were formed: 1) skills for all professionals; 2) some skills more than others; 3) yes, they need all the skills; 4) skills to become better practitioners; 5) usefulness to practitioners; 6) communication with public; 7) they're [the public] are depending on us; and 8) the future

***Keywords*:**

public health; informatics training; workforce development; needs assessment; education

Introduction

Unlike medical informatics, public health informatics is not a well established field. [1] Because medical informatics developed into a field first, much of the literature in American journals regarding informatics reflects the 'medical model.' There is little literature that reflects the 'public health model.'[2] Often medicine and public health are thought of as interrelated.

Medical professionals think of public health as a subspecialty of medicines and public health professionals view medicine as an arm of public health. [3] Each point of view has merit. Since managed care has encouraged clinicians to emphasize prevention and to examine the use of population-based outreach services, the differences between the two are becoming less evident. However, the two have different primary foci; the patient for medicine, and the community for public health. Also the role of information, core functions, data and information sources are more wide-ranging in public health practice. [1,2,3,4]

The role of information in public health

The foundation of public health is information. The 1996 World Health Report cites the continuing need to "disseminate health information widely, in the shape of epidemiological and statistical data, reports, guidelines, training modules and periodicals." [5] In practice, all the core public health disciplines – epidemiology, biostatistics, behavioral sciences, environmental health, and policy and administration – are supported by information and information technologies. Furthermore, according to the Institute of Medicine (IOM) the substance of public health is organized community efforts aimed the prevention of disease and promotion of health. It links many diverse disciplines and rests upon the scientific core of epidemiology. [2] Since epidemiology depends on the effective collection, analysis, interpretation, and dissemination of information, information is the foundation of public health.

Core functions of public health

In the IOM model, three core functions are necessary for sound public health practice: assessment, policy development, and assurance. A fourth core function, continuous evaluation, links the three core functions. [2] Each of these core functions is data and information driven.

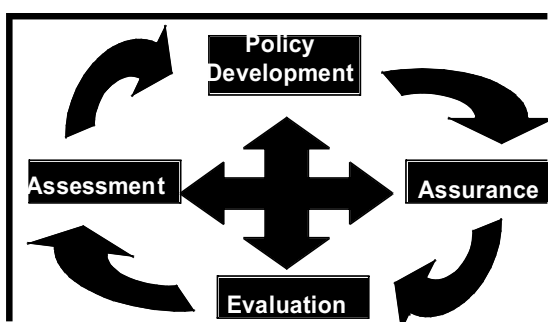


Figure 1 – Core Functions of Public Health

Assessment depends primarily on the core public health disciplines of epidemiology and biostatistics. [4] It consists of activities such as: surveillance; identifying needs; analyzing problems; collecting and interpreting data; case-finding; monitoring and forecasting trends; research and developing outcome indicators.

Policy development is a process by which public health professionals formulate, advocate and implement the appropriate action determined from the assessment of the situation. Policy development includes: decision-making;

mediating opposing views; negotiating common goals; developing programmatic goals and objectives; and allocating and mobilizing resources.

Assurance is the process that examines the quality of the situation, or impact of the policy or program implemented. Assurance activities include: determining the necessary services are provided to meet stated goals; regulating services and products; maintaining accountability and progress reports; and testing outcome indicators.

Evaluation is the ongoing assessment of the activities and processes. Evaluation data are gathered through a wide variety of sources.

Public health data sources

The public health core functions are performed in a data and information enriched environment. Data are collected through methods such as surveillance, surveys, and clinical assessments. The types of data generated include: demographic; mortality; morbidity; laboratory; physical assessments; health resources; health resource utilization; natality; disease occurrence; environmental assessments; and, public opinion data. There are as many sources for public health data as there are situations for assessment.

Public health information sources

Public health information sources also are unlimited. [6] Two common sources of categorized, meaningful data are Vital and Health Statistics and the National Health Interview Survey. Other more complete information sources include: the Morbidity and Mortality Weekly Report (MMWR); health department reports; health professions journals; newspapers; and magazines. Informal public health information sources such as phone calls, faxes, email, or meeting notes are a rich source of data and information about individuals and groups dealing with a specific public health situation.

Importance of public health information sources

Even with all the data and information sources available, public health professionals have difficulties obtaining the data and information they need. [5] Public health data and information sources lack standardization in organization, nomenclature and electronic transmission. Of the data are fragmented due to redundant collection or episodic data collection. Most notably, data and information is not always up-to-date. Timeliness of data is of importance. Nearly all public health actions depend on what is presently occurring in regards to a particular public health problem in a community.

If public health as a field is to become more effective, public health professionals need timely, quality information, better ways to communicate data and information, and betters tools to analyze new information. [1] Innovative methods of storing, organizing and disseminating the millions of pieces of data gathered during the core functions of public health need to be researched and tested. [6]

Because information is the major component of public health, and there is a need for public health professionals to know how to access, analyze and disseminate data and information. This study was designed to examine public health academician's belief in the importance of informatics for public health practitioners and the need to have informatics training in schools and graduate programs of public health.

Methods

This descriptive, qualitative study included 28 accredited schools of public health, 11 graduate programs in community health and 20 graduate programs in community health/preventive medicine which offer a MSPH or MPH degree within the United States.

A letter was sent to each school and program describing the study. A follow-up telephone call reached academicians at 21 public health schools, 14 graduate programs in community health/preventive medicine and 5 graduate programs in community health for a total population of 40 public health academicians. During this contact, one-on-one telephone interviews were scheduled.

Table 1. – Primary Contact in Study (N=40)

Number	Title
3	Dean
10	Associate Dean
1	Assistant Dean for Student Affairs
5	Professor
1	Department Head
4	Department Chair
5	Director of MPH Program
1	Associate Director of MPH Program
8	MPH Program Coordinators/Managers
1	Director of Student Services
1	Director of Distance Education

Data were analyzed from the interview surveys. The qualitative data analysis was conducted using a note-based analysis and open coding technique. [Note-based analysis involves immediately summarizing the notes take during the interview. Open coding describes the process of breaking down, examining, conceptualizing and categorizing the data.] Categories are the classification of concepts based on the comparison of one concept against another and those that appear to have similar characteristics are grouped together. [7]

Results

When asked "do you believe that informatics or informatics-related competencies are important for public health practitioners?" – 100% of the contacts said yes; several were emphatic. When asked why they believed informatics was important a variety of reasons emerged. Analysis of the comments revealed some interesting groupings.

Skills for all professionals

One subgroup of comments indicated these skills are necessary for any person to function in a professional capacity. This *skills for all professionals* subgroup had a global view of the need for informatics. The terms "computer age" and 'computerized society" were used to describe their belief; as one contact stated: "…you can't get along in the real world without these skills today."

Some skills more than others

Another subgroup of comments stated that some skills in informatics are needed by public health practitioners, but not all. This *some skills more than others* subgroup believe database use and development skills, information access skills, data and information analysis skills and an understanding of how communication technologies work is what should be taught. As one member of this subgroup succinctly stated: "…but not every practitioner needs to know all the details."

Yes, they do need all the skills

On the opposite end of the spectrum, the *yes, they do need all the skills* subgroup believes "there is a clear need to have expertise in informatics and the students need to understand the issues as well as the skills and knowledge." They believe the "shift to online based information is here to stay and [they] don't see any way to function in the public health arena without access to information technology sources." To teach these skills 'far outweigh the costs of the training and we need to support and train public health practitioners in this area."

However, "technology has gotten sophisticated so very quickly and those teaching need to understand the whole field of informatics and what it has to offer to public health. The faculty see only a narrow bit that is in their specialty area." In looking towards the future, they believe "we're not really preparing our public health students for the future" and "we think that it is criminal to graduate students without these skills."

Skills to become better practitioners

This subgroup believes informatics provides the tools to move the field of public health forward and there is a need for informatics *skills to become better practitioners*. They foresee that knowing when to use the "appropriate tools in each situation" will allow practitioners to be "more effective and cost-effective" in the practice of public health. Because this increased ability to be more effective, "informatics needs to become a basic professional skill." One person stated: "There is a clear need within public health to be able to use information technology."

Usefulness to practitioners

In this subgroup informatics *usefulness to public health practitioners* comments were categorized into different types: *information access*, *remote access,* and *cost-effectiveness*.

Widespread understanding of technology and information management concepts will "improve information access by practitioners" and allow practitioners to collaborate to "provide better services in their state or region." As one contact commented: "there is a lot of information to access and to be able to provide information in an intelligent way for others to find and use in practice, is a valuable need for public health."

Remote access was commented upon within this subgroup. In rural areas public health practitioners have limited access to information sources. The contacts believe information technology would greatly improve the practice of public health in these regions.

The cost effectiveness concept was stated by contacts in there belief that information technology has the ability to provide, "the latest developments in public health and has a lower cost and an easier accessibility than paper."

Communication with public

Another subgroup of comments characterizes the importance of informatics and information technology, the Web in particular, as methods to improve communication with the public. They believe it is "a vehicle by which we can communicate with the general public and with each other." They see expertise with information technologies necessary for public health practitioners to be able to "disseminate public health related information to the community and practitioners who want to implement intervention and prevention programs." A caveat to this improved communication is the frustration caused by lack of hardware in some populations and "there is a potential for creating a technology elite."

They're depending on us

An extension of the *communication with public* subgroup believes public health practitioners should take a leadership role in placing accurate health information on the web. The *they're depending on us* subgroup believes: "the world expects to find health information on the web"; the world expects that the health information on the Internet to be accurate"; and "public health has an obligation to provide accurate information [on the web]."

Future changes in practice

This subgroup believes the *future changes in practice* and information technology are inseparable. They believe graduates from schools and programs in public health "won't be employed in traditional public health programs, but in managed care organizations and the technology will become increasingly important in that arena," and "with managed care, from my occupational medicine point of view, data management is becoming more and more important to primary care." They also stated: "public health is moving away from the clinical services toward

the core areas of public health and within the core areas information is essential. For example, we need to maintain and access immunization registries to be able to accurately monitor health and well being of children and have the people with the skills to maintain them," and "in public health state departments [we're] moving from the delivery of services to the evaluation of programs which will emphasize the need for them [public health workforce] to have informatics skills and knowledge.´ This subgroups was very ardent in their beliefs. One contact summed up the thoughts persuasively by stating: *this is not a trend, but a modus operandi for the future*.

Conclusions

All forty contacts, regardless of their academic affiliation, stated they believed informatics and informatics related concepts were important to the practice of public health. They understand the importance of timely, accurate, and quality information. Yet, they seem to believe the information will automatically appear into information systems. To have the quality of data and information needed to practice public health effectively, the data must be gathered and organized in a standardized manner, and made available in an easily accessible form within information systems. The ability for public health practitioners to develop systems to ensure accessible, quality information should be provided through public health informatics training.

The emphasis on the information component is important, but there is value in understanding the technology. The distinct advantage of training in public health informatics is: The practitioner would not only have knowledge of what the data was needed for; they would be able to determine what technologies would be best suited to deliver the data and information expediently in the most useful format; and they would have the skills and knowledge to best organize the data or information to facilitate communication to other public health practitioners.

There is a need to develop information systems and technology to improve public health practice. Schools and graduate programs of public health are slowly developing courses in public health informatics to meet this need. [8] They believe informatics is important to the future of public health.

References

[1] Friede A, Blum HL, and McDonald M. Public health informatics: how information-age technology can strengthen health. Ann Rev of Pub Hlth 1995:16:239-52.

[2] Institute of Medicine (IOM). The future of public health. Washington DC: National Academy Press, 1988.

[3] Voelker R. Medicine and public health join forces. JAMA 1997: 277(20) : 1579.

[4] Afifi AA and Breslow L. The maturing paradigm of public health. Ann Rev Pub Hlth 1994: 15: 223-35.

[5] WHO. The world health report, 1996: fighting disease, fostering development. WHO 1997: 1-12.

[6] Friede A and O'Carroll PW. Public health informatics. In: Wallace, RB and Doebbeling BN eds. Maxcy-Rosenau-Last. Public Health and Preventive Medicine. 14th ed: Stamford Connecticut: Appleton & Lange 1998; pp 59-66.

[7] Strauss A and Corbin J. Basics of qualitative research: grounded theory procedures and techniques.

[8] Richards J. Informatics Training in Public Health. Masters of Public Health Thesis. The University of Texas School of Public Health. 1999.

Acknowledgments

Thank you to the generous support of the University of Texas system and the participants in this study.

*This study was completed at the University of Texas – this paper reflects the opinion of the author and not necessarily those of the Centers for Disease Control and Prevention.

Address for correspondence

Janise Richards, MS, MPH, PhD, National Center for Public Health Informatics, Centers for Disease Control and Prevention, 1600 Clifton Road, M/S E85, Atlanta, GA 30333

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Educating Medical Students as Competent Users of Health Information Technologies: The MSOP Data

Julie J. McGowan[a,b], Morgan Passiment[c], Helene M. Hoffman[d]

[a]*School of Medicine, Indiana University, Indiana, USA;* [b]*Regenstrief Institute, Inc., Indiana, USA*
[c]*Association of American Medical Colleges., Washington, D.C., USA*
[d]*School of Medicine, University of California San Diego, California, USA*

Abstract

As more health information technologies become part of the health care environment, the need for physicians with medical informatics competencies is growing. In 2006, a survey was created to determine the degree to which the Association of American Medical College's Medical School Objectives Project (MSOP) medical informatics competencies had been incorporated into medical school curricula in the United States. Methods: a web-based tool was used to create the survey; medical education deans or their designees were requested to complete the survey. Analysis focused on the clinician, researcher, and manager roles of physicians. Results: Seventy usable surveys were returned. Many of the objectives were stated in the schools' respective curricula and the competencies were being evaluated. However, only a few schools taught and assessed the medical informatics objectives that required interaction with health information. Conclusion: To insure that physicians have the knowledge, skills, and attitudes to effectively and efficiently interact with today's health information technologies, more medical informatics concepts need to be included and assessed in all undergraduate medical education curricula in the United States.

Keywords:

education, medical, undergraduate; medical informatics; hospital information systems; decision support systems, clinical

Introduction

Within the next decade, a large majority of hospitals and health care centers in developed nations and many in developing nations will have electronic health records and other forms of health information technology. Physicians will be expected to use these tools to improve patient safety, enhance the quality of care, and reduce costs. This expectation requires that physicians be trained, not as medical informaticians but as knowledgeable users of the health technology tools. However, most education in medical or health informatics has focused on the knowledge and skills needed by informaticians rather than health care professionals.

Recently in the United States, the President authorized the creation of the first Office of the National Coordinator for Health Information Technology. Several legislative initiatives were undertaken to promote the use of information technology within healthcare to improve process, quality and safety, thereby improving the health of our citizens. The promise of widespread adoption of electronic health records with the concomitant capabilities of provider order entry, decision support, and data mining for clinical research, as well as quality and safety evaluations, is about to become a reality. However, significant questions exist as to whether or not physicians will have the competencies necessary to effectively use these systems to achieve the goals outlined by the President and legislature.

Europe and Canada have long been leaders in the training of informatics-facile health care providers. The work of the European Centre for Medical Informatics, Statistics and Epidemiology (EuroMISE) has provided an early framework for such education in Europe.[1] The International Partnership for Health Informatics Education is in part an outgrowth of the earlier efforts and, in an environment of increasing globalization, emphasizes the need for international components in informatics education.[2]

Canada was also an early leader in medical informatics education and took a different but equally effective approach by integrating applied medical informatics into the undergraduate medical curricula.[3] However, such education must evolve with the changing technologies and the demand for more and more health care professionals to become information literate has resulted in an evaluation of current practices with more emphasis being given to emerging trends in both informatics and health.[4]

Other nations are beginning to recognize the need for more informatics training in the health professions.[5-7] In an attempt to address these very real issues, the International Medical Informatics Association developed recommendations on education in health and medical informatics.[8] These recommendations are the initial step in developing the educational framework necessary to insure that students possess appropriate qualifications to work in an information technology intensive health care environment.[9]

Leaders in medical informatics in these countries and others are calling for more targeted educational programs to insure that the systems being implemented will have physicians trained to use them.[10] However, the integration of such training into health professions curricula has been difficult at best and quite slow to develop.

Need for such training was beginning to be recognized in the United States in the 1980s with several calls from major organizations to prepare physicians for a future in an automated health care environment by integrating the necessary skills into the educational process.[11-13] However, little was realized in the form of concrete programs from these early inducements.

Understanding the potential impact of the growing interest in health information technology on the practice of medicine, and trying to take a more proactive stance in insuring that undergraduate medical students had a firm grounding in the knowledge, skills and attitudes necessary to become technologically savvy health care providers of the future, the Association of American Medical Colleges in 1998 convened an expert panel to develop educational objectives to satisfy this goal. The medical informatics panel of the Medical School Objectives Project (MSOP) II identified five medical informatics relevant roles played by physicians – lifelong learner, clinician, educator-communicator, researcher, and manager. The recommendations for educational content were developed within this framework and published in 1999.[14]

In part because of the increasing interest on the part of the government in facilitating widespread adoption of health information technology, in part because of the dearth of articles published about new educational programs in medical informatics in undergraduate medical curricula, and in part because of a growing need for information literate physicians, a small group of the educational leadership within the Group on Information Resources of the Association of American Medical Colleges surveyed and analyzed the responses of the 127 United States medical schools to determine whether or not they had implemented the MSOP medical informatics educational objectives and, if so, to what extent were the implemented.

Methods

An initial request to participate in the survey was sent to the respective deans of medical education at the 143 discrete medical schools in the United States and Canada. The deans were asked to either respond to the survey or refer it to someone who was knowledgeable about medical informatics content in the curriculum. The Web-based survey asked participants to respond to questions formulated directly from the MSOP II medical informatics educational objectives. These questions were grouped by the physician role with sub-groupings around concepts.

An initial question addressed whether or not the respondent was familiar with the MSOP Medical Informatics educational objectives. The subsequent questions asked the respondent whether or not each of the objective concepts was taught, had stated objectives, and was assessed. At the end of each of the five role divisions the respondent was asked to indicate who taught the concepts and how the concepts were assessed.

While virtually all of the respondents indicated they were familiar with the MSOP medical informatics educational objectives, the responses differed widely in regards to teaching, stated objectives, and assessment. In following up with a number of the participants about responses, it became apparent that many thought the medical informatics content was being taught as an integrated component of the clinical years. However others from the same institutions, many with long standing clinical information systems, stated that their medical students were exposed to these systems but did not have formal training or experiential learning with these systems.

Because of these discrepancies, a second survey was developed that limited responses to stated objectives and assessment because of the belief that having a stated objective would result in some educational action and would eliminate the possibility of someone assuming rather than knowing that the concepts were being taught.

The request to participate in the survey was again sent to the deans of medical education unless there was a different respondent on the first survey. The second survey was also Web-based and a request for participation was made in early 2006, almost a year after the first survey. Interestingly, individual school responses showed little change, however, several additional schools indicated establishing objectives.

Because the attributes for being facile with health information technology in the today's health care environment focused on three of the five physician roles, the responses for Life-long Learner and Educator-Communicator were not considered for this study. In addition, while data was collected on Canadian medical schools, because of their early embracing of the need to teach medical informatics in undergraduate medical education, only the responses from United States Medical Schools have been evaluated.

Results

Seventy usable surveys were "virtually" returned. Ninety-six percent of the respondents were familiar with the MSOP medical informatics educational objectives and eighty-eight percent indicated that there had been an overall strategy to integrate medical informatics objectives into the curriculum. However, the results of the specific competencies did not support this.

Clinician

Within the sub-group of effective use of clinical information systems, 60% of the respondents indicated that they had a stated objective on retrieving patient-specific information from a clinical information system and 49% assessed the competency. Forty-four percent had a stated objective on displaying selected subsets of information available about a given patient and 36% assess the competency. Forty-six percent had a stated objective about recording specific findings about a patient in a clinical

information system while 47% assessed the competency. Forty-six percent had a stated objective on recording orders (CPOE) directing the further care of the patient and 36% assessed the competency.

The sub-group of interpreting laboratory tests scored higher. Seventy percent of the respondents had a stated objective about recognizing the knowledge limitations of standard laboratory measurements and 66% assessed the competency. Seventy-seven percent had a stated objective about demonstrating the ability to integrate clinical and laboratory findings while 86% assessed the competency.

Within the sub-group of incorporating uncertainty explicitly into clinical decision making, fifty-seven percent of the respondents had a stated objective on demonstrating the ability to quantify and communicate the degree of certainty associated with specific items of scientific and clinical information and 50% assessed the competency. Forty-six percent had a stated competency on demonstrating the ability to identify and locate when possible the crucial pieces of missing clinical information and determine when it is appropriate to act on incomplete information and 40% assessed the competency. Sixty-three percent had a stated objective on demonstrating the ability to integrate verbal and statistical sources of medical knowledge with the facts of a specific clinical case and 61% assessed the competency.

Within the critical use of decision support tools sub-group, sixty-nine percent of the respondents had a stated objective on using textbooks and journal articles and 67% assessed the competency. Thirty percent had a stated objective on using diagnostic expert systems and fourteen percent assessed the competency. Twenty-three percent had a stated objective on using advisories or alerts issued from a computer based records and fourteen percent assessed the competency.

In responding to a student's ability to formulate a treatment plan, fifty-seven percent of the respondents had a stated objective that students should demonstrate the ability to express the relative certainties of a differential diagnosis while 69% assessed the competency. Sixty-one percent had a stated objective on expressing the relative risks and benefits of outcomes and treatment options while 66% assessed the competency. Forty-six percent had a stated objective on taking action by balancing risks and benefits while 53% assessed the competency.

Within the sub-group of respecting patient (and physician) confidentiality, 76% of the respondents had a stated objective on demonstrating the knowledge of the legal, ethical and medical issues surrounding patient documentation including confidentiality and data security while 79% assess the competency. Thirty-three percent had a stated objective on demonstrating the ability to use security-directed features of an information system while 27% assessed the competency.

Researcher

The first of the researcher group deals specifically with the use of clinical information systems. Twenty-four percent of the respondents had a stated objective on determining a practice's case mix and 20% assessed the competency. Twenty-nine percent had a stated objective on determining the incidences of diagnoses in a practice and 26% assessed the competency. Forty percent had a stated objective on testing the efficacy of a new treatment and 33% assessed the competency. Fifty-six percent had a stated objective on formulating testable hypotheses and 50% assessed the competency. Fifty-one percent had a stated objective on collecting, organizing, and interpreting data while 53% assessed the competencies.

Within the sub-group about determining what data exist relative to a clinical question or formal hypothesis, seventy-one percent of the respondents had a stated objective for demonstrating the ability to use information technology to locate existing data sources and 60% assessed the competency. Thirty-three percent had a stated object for demonstrating knowledge of data sources (including medical records claims and reimbursement information and online data) at one's own institution by identifying how these might be used to address a specific clinical question posed as research and 20% assessed the competency. Thirty-one percent of the respondents had a stated objective for demonstrating the ability to identify and locate existing data sets no maintained at one's own institution (e.g., national registry data) that might be used to address a specific clinical question posed as research and 16% assessed the competency.

For the sub-group executing a plan for data collection and organizing data for analysis, 24% of the respondents had a stated objective for selecting and appropriate computer database tool for collecting and organizing data and fourteen percent assessed the competency. Twenty-nine percent had a stated objective for properly representing data from a study in a form that is useful and supports computer-based analysis and sixteen percent assessed the competency.

Within the sub-group of analyzing, interpreting, and reporting findings, 23% of the respondents had a stated objective for selecting the appropriate computer software tools for analysis of data and ten percent assessed the competency. Thirty-one percent had a stated objective for using software to perform simple statistical analysis and portraying the results graphically and 23% assessed the competencies. Thirty-one percent had a stated objective for interpreting the reports of statistical software analysis and 27% assessed the competency.

Manager

There are three sub-groups within the Manager grouping. The first of these is the appreciation of the role of information technology in relation to managing the cost of medical care and its impact on individuals and society. Twenty-three percent of the respondents had a stated objective on using on-line sources of health care financing information and eleven percent assessed the competency. Thirty-nine percent had a stated objective on continuous quality improvement and process management and twenty percent assessed the competency. Twenty-four percent had a stated

objective on how information technology can be used to develop, implement and monitor compliance with clinical pathways and other forms of patient care protocols and eleven percent assessed the competencies. Thirty-three percent had a stated objective on how clinical information in the aggregated can be used to determine health care services planning for populations and 23% assessed the competency.

Within the sub-group of formulating and making decisions for individuals and groups, 55% of the respondents had a stated objective on demonstrating knowledge of cost/benefit issues in health care and 29% assessed the competency. Fourteen percent had a stated objective on using a decision-analysis package and seven percent assessed the competency. Thirteen percent had a stated objective on using software utilities assessing patients and six percent assessed the competency. Thirty-nine percent had a stated objective on incorporating economic and cost perspectives into decision making and 23% assessed the competency.

The last sub-group dealt with working effectively as an individual in inter-professional groups and as a member of a complex health care system. Nineteen percent of the respondents had a stated objective on using electronic personal and clinical scheduling systems and nine percent assessed the competency. Twenty-one percent had a stated objective on archiving and organizing digital information of personal and clinical import and fourteen percent assessed the competencies. Twenty-four percent had a stated objective on demonstrating knowledge of online resources for legislation, political advocacy, and local health care policy setting and six percent assessed the competency.

General questions

In all three of the physician role groupings, the content was taught generally through embedding it in core course. A few schools had an elective course in medical informatics and fewer still had a core course in medical informatics. Because the primary mode of teaching was through integration with other content, almost all of the assessment of competencies was done as part of a general educational evaluation schema. However, several schools had tests specific to medical informatics or used these in conjunction with the general assessment methodologies.

Discussion

The medical informatics educational objections presented by the MSOP expert panel were developed around the concept of information discovery and not predicated on computer literacy. For this reason, a number of the competencies can be taught without use of a computer. Examples of this are found in the interpretation of laboratory tests and the ability to formulate a treatment plan.

There were a total of 41 questions in the clinician, researcher, and manager role groups. Of those, 27 required interaction with a clinical information system or some ancillary system containing patient information. Eleven questions involved educational objectives that could be met without such interaction. Three questions related specifically to the competencies within the life-long learner role group but were also closely linked to clinician and researcher information management.

Of the five roles, the greatest number of medical school having stated objectives and competency assessments was found in the life-long learner role. This corresponds to the increase in teaching evidence-based medicine and the greater involvement of libraries for development of knowledge-based searching capabilities. For this reason, the life-long learning correlates, although requiring the use of computers to find information, were grouped separately.

In analyzing the responses by question type, less than a third (30.7%) of the medical school respondents had stated objectives for the 27 questions requiring use of computer systems and only slightly more than a fifth (21.1%) assessed competencies. There was one exception. Sixty percent of the respondents did have a stated objective about retrieval of patient-specific data from a clinical information system and 49% assessed the competency.

Of the three life-long learner correlated questions, approximately two thirds (67.7%) of the medical school respondents had stated objectives and slightly less (62.7%) assessed the competencies. Of those questions that did not require interaction with a computer system, over half (58.6% and 56.7% respectively) of the medical school respondents had both stated objectives and assessment of competencies.

In looking at the raw data and comparing the assessment to the stated objectives in all three of the physician role groups, there were 28 instances in which competencies were assessed within seven sub-groups without having stated objectives. These were virtually all in the clinician role and fell primarily under the non-computer based questions. A possible explanation is that the concept might have been considered too granular to include as a stated objective while it was included as part of a clinical evaluation schema.

Conclusion

Seventy of 127 surveys assessing the degree to which the MSOP medical informatics educational objectives have been incorporated into undergraduate medical curricula in the United Stated were completed. An analysis of these found that while many of the medical informatics concepts relevant to the clinician, research and manager roles were being addressed in the curricula, when broken down by those concepts that required health information technology interaction, only a few schools had stated objectives and fewer assessed the competencies.

The survey respondents were self-selected, and anecdotal information suggests that many who did not complete the surveys chose not to do so because they had little or no medical informatics in their curricula. Also, while these objectives are valid today, as HIT systems evolve and become more integrated into the health care system, the objectives also need to evolve. Some progress has been

made but much more needs to be accomplished to insure that physicians will be able to efficiently and effectively use the health information technology being installed in hospitals and health centers.

Acknowledgements

The authors gratefully acknowledge the support of the Group on Information Resources of the Association of American Medical Colleges for facilitating the survey distribution.

References

[1] Zvarova J, Engelbrecht R, and van Bemmel JH. Education and training in medical informatics, statistics and epidemiology in EuroMISE. Int J Med Inf 1997: 45: 3-8.

[2] Jaspers MW, Gardner RM, Gatewood LC, Haux R, Schmidt D and Wetter T. The International Partnership for Health Informatics Education: lessons learned from six years of experience. Methods Inf Med 2005: 44(1): 25-31.

[3] Kaufman DM and Jennett PA. Preparing our future physicians: integrating medical informatics into the undergraduate medical education curriculum. Studies Health Tech & Info 1997: 39: 543-6.

[4] Moehr JR. Where to in the next ten years of health informatics education? Methods Inf Med 2006: 45(3): 283-7.

[5] de la Calle G, Maojo V and Benito M. The INFOBIOMED Network of Excellence: facilitating training and mobility in biomedical informatics in Europe. Stud Health Technol Inform 2006: 124: 893-8.

[6] Sarbadhikari SN. Basic medical science education must include medical informatics. Indian J Physiol Pharmacol 2004: 48(4): 395-408.

[7] Hasman A and Haux R. Curricula in medical informatics. Stud Health Technol Inform 2004: 109:63-74.

[8] Haux R. Health information systems – past. Present, future. Int J Med Info. 2006: 75(3-4): 268-81.

[9] Muller S (Chair). Physicians for the twenty-first century: report on the project panel on the general professional education of physicians and college preparation for medicine. J Med Educ 1984: 59(11pt2): 1-208.

[10] Myers JD (Chair). Medical education in the information age. Proceedings of a symposium on medical informatics. 1986: Washington, DC: Association of American Medical Colleges.

[11] Association of American Medical Colleges. Advisory Panel on Assessing Change in Medical Education – The Road to Implementation. ACME – TRI Report: Educating Medical Students. Assessing Change in Medicine Education. The Road to Implementation. 1992: Washington, DC: Association of American Medical Colleges.

[12] Association of American Medical Colleges. Contemporary issues in medicine – medical informatics and population health: report II of the Medical School Objectives Project. Acad Med 1999: 74(2): 130-41.

Address for Correspondence

Julie J. McGowan, Ph.D., FACMI
Indiana University School of Medicine
975 W. Walnut Street IB-310
Indianapois, IN 46202-5121 USA

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Establishing A National Resource: A Health Informatics Collection To Maintain The Legacy of Health Informatics Development

Beverley Ellis[a], Jean Roberts[a], Helen Cooper[b]

[a] *Health InformaticsU nit, Lancashire School of Health and Post-Graduate Medicine, UCLAN, Preston UK*
[b] *Health Team, Library and Learning Resource Service, University of Central Lancashire, UK*

Abstract

This case study report of the establishment of a national repository of multi-media materials describes the creation process, the challenges faced in putting it into operation and the opportunities for the future. The initial resource has been incorporated under standard library and knowledge management practices. A collaborative action research method was used with active experts in the domain to determine the requirements and priorities for further development. The National Health Informatics Collection (NatHIC) is now accessible and the further issues are being addressed by inclusion in future University and NHS strategic plans. Ultimately the Collection will link with other facilities that contribute to the description and maintenance of effective informatics in support of health globally. The issues raised about the National Health Informatics Collection as established in the UK have resonance with the challenges of capturing the overall historic development of an emerging discipline in any country.

Keywords:

health Informatics, legacy, knowledge management, access

Introduction

Existing, lost and retained multi-media materials relating to topics in informatics to support health care represent a challenge in the context of the complexity of the current Health Informatics (HI) landscape. Absence of materials is most frequently felt by students and researchers, although the HI reference content also directly frames current operational developments. Looking at past projects and processes, and their longevity, success and failure may give pragmatic indicators for the level of acceptance or reluctance to embrace current development. If literature generated in the initial, basic and developmental phases of the projects are still available for reference, it will be possible to identify milestones and situations that can explain current entrenched positions and perhaps suggest catalysts to success that were missed on the first iteration.

The closure of the NHS in England Information Authority and the realignment of its roles and responsibilities, predominantly under the NHS Connecting for Health (NHSCFH) agency National Programme for IT (NPfIT) [1], made its reference resource library surplus to requirements. Members of the Health Informatics Unit at the University of Central Lancashire School of Health and Post-graduate Medicine (UCLAN) [2] had, with other academic colleagues, been considering how to sustain the legacy of HI. UCLAN was successful in bidding to house the residual resource; this is mainly due to its track record in managing and making available other specialist collections, such as the Livesey Collection relating to the Temperance Movement in the early 1800s, and the media collection of the Sport England organisation.

The National Health Informatics Collection (NatHIC) entity has been launched as a Collection rather than an Archive in order to confirm the commitment to maintain the heritage as a living and growing source. Personal donors who are, or were, key players in the domain have added to the initial materials, and continue to do so.

In order to keep a live Collection, the NatHIC objectives encompass making links to valuable knowledge facilities established elsewhere and to bodies who make available organisational resources on a rapidly changing basis. The linkages are in addition to managing physical resources on-site at UCLAN. The Collection is being made available for the benefit of all interested parties, not just those with close links to our university. NatHIC is intended to be available by reference, and where possible, loan of its content. NatHIC aims to act as a hub, maintaining physical multi-media resources, signposting other collections and materials, making linkages with other facilities and providers and providing guidance on identifying sources to those interested in informatics to support the health care domain. Users will not just be within the NHS but will also be from other care delivery agencies, their (commercial) solutions and service providers, and academia in all of the home countries and beyond. The desirable content will span both the historical legacy of HI and its contemporaneous / future materials, from a UK and wider perspective.

Method of development

The initial establishment of a repository of materials relating to the development of health informatics across the UK was first considered in the 1990s by members of CHIRAD,

Selected for best paper award.

a research institute in Winchester [4] whose members had had long and active participation in both teaching and learning and in the development of domain through the British Computer Society professional learned society and its subsidiary body, now called the Health Informatics Forum [5]. Despite various funding proposals the resources were not at that time available. However, the situation had changed sufficiently that in 2006 when substantial materials for the Collection became available, it was possible to put ideas into action at UCLAN. This Case Report paper describes the current situation, and discusses issues for future development in terms of content and accessibility for the domain internationally.

Access

Logistically, all users need to be recognized as authorized to use the Collection if they wish to loan any content. They must initially be registered as Distance (Associate Staff) Users of the UCLAN Library. This is a process already deployed in order to facilitate and enable generic library access by local NHS employees in the North West of England. It requires only completion of a form and provision of a photographic image to be used as authorisation if the holder ever wishes to physically enter the library on campus. UCLAN strives for equity of access for all through being cognisant of disability equity criteria including SENDA [6] Disability legislation.

NatHIC is initially intended to encompass any materials that impacts upon the HI developments across all the four UK home countries. Its principles and the lessons learnt from its establishment will perhaps be of use to other countries starting to identify similar omissions in their HI history. It is hoped that NatHIC will eventually interface to other resources elsewhere giving a global picture of HI developments over time.

Limitations

The current processing of donated will only capture legacy material that will set health informatics developments into context. UCLAN are putting into place mechanisms to include new works donated by publishers, authors and organisations creating informative resources over time.

Commentators have questioned the value of a back catalog rather than solely a comprehensive, contemporaneous resource. There are many significant instances where prior work has a distinct impact on current thinking. One particular instantiation of the continuing relevance of past concepts is a data quality principle of the early 1980s, where healthcare practitioners lead by the late Dame Edith Korner stated that a healthcare organisation should only (be required to) collect that information 'without which it is not feasible to be deemed to manage effectively' [7]. Thus data had to be collected at source, where any errors would be more readily identified and corrected; and that data had to be owned and have a purpose at that level in order to ensure the highest quality was maintained. At the time (1983) this reduced the number of resource wasting ad hoc studies that were previously carried out to answer one-off queries from government and strategic bodies. That principle can be seen reflected in the NPfIT 'Do once and Share' initiative but many younger HI professionals do not grasp that the underlying principles was defined over 25 years ago!

Others suggested that the current classification of cited publications was sufficient to afford a useful view of HI over time. However, as Machan [8] identified at the MIE2006 congress in Maastricht , analysis of the cited literature has a considerable omission in its knowledge capital – that of the negative findings and problem issues that are not formally published, but may have come to light in 'grey literature', trade papers and transient materials, such as newspapers, blogs and electronic commentaries.

As the Collection enlarges, there is a risk that space will become a premium. At that time, a usage review may be necessary to refine the most frequently requested sources to a particular geographic, technological, health or time focus, yet to be identified. Current library issue tracking systems can contribute to profiling what materials can/should be culled and can inform debate about what criteria should be used for content management.

Decisions relating to the scope, content, eligibility of materials for the Collection cannot be left to knowledge managers, library professionals or a domain experts with a close personal interest, (for example-ourselves) in maintaining the Collection, so an Editorial Management Committee with a range of areas of expertise is being established to reflect the extensive inclusive nature of the HI domain.

Determining ongoing functionality

Immediately after the launch of the Collection, a collaborative action research methodology was deployed in a Master Class facilitated by the NHS HI Faculty [9]. The session reflected on the basic Collection and its functionality and identified areas for further enhancement. These suggestions will be addressed over time in order to achieve as high quality resource as is feasible, given current technologies, funding and resources. The findings from the session were subsequently written up and validated through wider dissemination and feedback through the Faculty web area.

Various questions were identified, relating to establishing and maintaining resources for the benefit of all interested parties, they are explored further in this paper.

- What are the core resources that should be preserved
- What is the longevity of formal publications in practice
- How is it best to make available access to historic and future materials relating to HI
- Can a mechanism be developed that will accommodate self-input to the repository and yet avoid 'vanity' publication
- What target materials should be included in (or accessible through) the remit of NatHIC

Why is a comprehensive resource necessary

Health Informatics is still an emerging discipline, even after over thirty years [10] and as such is frequently seen as

an operational adjunct and a fragmented research domain. There is evidence that students, researchers and commentators on HI developments over time had difficulty finding historic materials to:

- set the context for current work
- explain the strategic direction of HI over time
- provide details of major national and local HI initiatives that were completed or had refocused for subsequent phases

In addition, there was anecdotal evidence, strongly articulated by professionals in the field that current initiatives were, or were likely to, repeat the mistakes and not able to avoid pitfalls that had been negotiated previously [11] if they did not take into account the way such challenges were addressed in the past.

Core content

It is not logical for NatHIC to deploy scarce resources to replicate / duplicate established sources of relevant material unless they are thought to be at risk. NatHIC is in the process of exploring links and collaborative working with repositories such as US National Library of Medicine, UK National Library of Health, and Department of Health with a view to providing secure gateways to their resources from NatHIC.

In order to be readily accessible and searchable, NatHIC will require ('a smart') classification and categorisation of its content. In view of the eclectic nature and wide range of that content, which cuts across traditional library resources, further work is ongoing exploring the suitability of existing ontologies and classifications.

Currently, the inclusion of materials is organic, decisions being based predominantly on advice about the seminal nature of thematic materials; but in due course a strategic acquisition plan for materials that the community / peer group see as important will be necessary.

The materials, nature and the links to NatHIC confound current definitions so the current structured strategies to facilitate effective searching are incomplete. For example – promotional literature aimed at a lay audience that is currently used to explain a proffered guarantee relating to care records handling will no doubt be varied because of current concerns [12]. It will be necessary to retain copies of the original documents to contextualise the emotions surrounding 'opting in or opting out' of the formal shared records plans for the 'National Spine' [1].

It was suggested that NatHIC also explored the management and retention of Case Study reports describing both completed assignments and projects still in progress in addition to issues papers created by various sources over time. There are a number of active groups in the field already cataloging such information, however these are disparate and diffuse, predominantly managing only the status reports of projects they themselves funded or support (for example Department of Health Service Delivery and Organisation R&D Programme : www.sdo.lshtm.ac.uk/commissioned projects.htm or the European Commission Europa site : europa.eu). NatHIC therefore could bring together linkages to those sources to produce a full picture of research activity across the country, and by UK participants wherever the project is lead from.

Also felt to be useful for ad hoc reference and download were templates and schemas that previously were only made available with specific initiatives and not available at other times, including business case presentations, project initiation documents and audit schemas.

Format

The permissible media for submission of items for the Collection were felt to be less crucial than the content. However it was recognised that some media were made obsolete by changing technologies. It is the intention of UCLAN, over time, to explore the development of mechanisms to make best efforts to keep seminal material readily accessible. For example – degraded old documents can be made available in scanned electronic form; obsolete e-media can be re-versioned forward as required (such as VHS video formats being transferred to CD-ROM). The experiences of other locations, such as the British Library will be sought to frame this activity.

NatHIC 'material content' will consist of physical hard copy media, audio and video/CD-ROM in so far as are still operational, robust electronic documents where they are provided by a source organisation (and subsequently where necessary to preserve content from obsolescence) and also links to other repositories, sources of content.

Where possible, donors will be encouraged to provide electronic media and direct access to back catalogues.

It was identified that users of NatHIC may only have a requirement for selected materials from within a journal / report and that partial download facilities should be available wherever possible. This will require a level of pre-processing which is not yet in place.

Moderation

As stated previously, NatHIC is currently taking all materials offered to it, with the knowledge that some may be extraneous, duplicate or out of scope. If the interest in NatHIC continues to rise, content may rise indeterminately and there will be a need for moderation of 'copy' in order to preserve the legitimacy of the resource. This will need to be carried out without bias by domain experts and knowledge management staff in the light of a number of factors, such as:

- content value to the domain
- scope of the topics and themes deemed appropriate
- usage rates
- copyright and license criteria

Value to the domain may change over time and will need periodic review. This could result in some material being withdrawn as no longer of significant worth or retained upon being confirmed as seminal and necessary to remain in the Collection.

Categorisation of materials will include evaluation by some, as yet un-constituted body, as relevant, in-scope and worthy of inclusion. In the immediate future, all received material will be stored. It will then be reviewed and processed for long term retention when appropriate criteria are determined.

The volume of copies submitted may require management against their usage; some key documents like national audit reports may be on frequent access by many people. A policy of retention for best quality copies being retained is already in operation, subject to estimated usage.

The mid-term requirement for moderation will necessitate drawing up rule-based policies for relevance, inclusion / exclusion and retention / withdrawal of material to ensure the integrity of NatHIC and its objective to preserve history and protect the future. This cannot be developed in isolation, but will be framed by the incoming material, potential user requirements and overall scope, and agreed by a representative editorial board, as yet un-constituted. In addition, for materials given to NatHIC, standard library protocols and existing mechanisms for clearing / establishing copyright and usage permission will be validated to determine their appropriateness; these materials include public domain materials, in-house reports etc., personal publications (for example student dissertations) or materials previously published elsewhere.

Future functionality

Facilities for self-loading of relevant publications may be useful; both by leaders in the field and for first line exposure of student theses and scientific papers. However draft terms and conditions will require definition and consultation, and processes for review will need to be operational to avoid / limit vanity publishing of dubious quality and contribution to the domain.

Grey literature, currently un-cited, ill-defined and transient may present a valuable resource to complement existing publications and complete the funnel plot described by Machan and demonstrate full coverage of knowledge across the domain. Further research will be necessary to validate this premise. In the interim, grey literature will be pursued, collected and reviewed at a later date.

Commercial bookshops around the world provide a service to new readers that indicates what the views of previous readers were on a particular publication. These views are displayed with books for sale on the shelves, but one suggestion is that a similar electronic facility could be added to NatHIC to allow users to comment on, or recognize similar materials on a topic that is in the catalog.

NatHIC aims to be inclusive and act as a portal where other established sources already have Health Informatics materials that are relevant. This will require planning and negotiation and will happen incrementally.

Critical success factors for NatHIC

Initially, success would be measured by a reduction in the number of historic documents being 'lost' to the community. Subsequently, it may be possible to monitor 'unrequited hits' for requested materials as an indicator of whether NatHIC is meeting community needs. As a proxy, an increasing number of formal arrangements with key stakeholders to routinely provide updated material will indicate increasing success.

The NatHIC has a linked informal blog [12] that will capture data on suggestions for new material and links. Monitoring unrequited searches and feedback will be reviewed periodically, so that NatHIC can target areas of interest that are currently not fully satisfied.

Results

To date, by early 2007, the Collection has over 1,500 contributions within an overall university library catalog of over 608 thousand items, including over ten thousand journals in the medicine and social sciences thematic areas. The range of media includes both contemporary hard copy and CD/DVD material and 'at risk' ephemeral media – including VHS videos and degrading paper. The 'at risk' material is being catalogued and steps will be taken to migrate it to viable media to capture the content.

Amongst the material provided are items which track the sequential phases of strategic development of NHS computing in England, specific business specifications of departmental and clinical systems, research deliverables from the sixties onwards and guidance on audit and information governance in both the NHS and European contexts.

The earliest material so far is - Berne, Eric (1968) Games people play: the psychology of human relationships [National Health Informatics Collection] Harmondsworth : Penguin, 0140027688. It is our belief that earlier material exists and is still being catalogued and prepared for NatHIC. HI in the UK is identifiable from 1961 when the first in-house hospital system was established in Manchester, although arguably, the NHS Central Registry pre-dates this as used in World War II to ensure unique patient identification. NatHIC is used by international cohorts of UCLAN HI students on Foundation degree to Masters courses; nearly 50 interested external parties have requested registration as Associates of UCLAN in order to access the NatHIC. This low number will increase as the body of content increases and as we attend events and promote the service.

Conclusions

The establishment of a national collection of Health Informatics material has been long overdue in the UK. It brings together resources that previously were (thought) lost or destroyed or considered to be of transient value. The advent of innovative electronic media makes this task feasible and sustainable in a manner not previously possible, and makes the National HI Collection accessible to a broad range of potential users. Usage over time will indicate its success. At this point in time, historic content is being added to the collection on a regular basis and arrangements with publishers for new materials to sustain

the Collection as a living archive are underway. In addition, newly published material will be incorporated as and when publishers and authors lodge reference copies with the Collection. The lessons learnt by this development will have resonance with any researcher who has previously not been able to source a required text on any media.

Acknowledgements

The authors thank colleagues working in and for the NHS, now and in the past, for their generous contributions to the content and development of the National Health Informatics Collection.

References

[1] Connecting for Health National Programme for IT, NHS in England www.connectingforhealth.nhs.uk (accessed December 2006)

[2] Health Informatics Unit outline, Lancashire School of Health and Post-Graduate Medicine, University of Central Lancashire www.uclan.ac.uk/healthinformatics (accessed March 2007)

[3] Livesey Temperance Society Collection, www.uclan.ac.uk/library/usersupport/lrs/collection/livesey/index.htm (accessed March 2007)

[4] Centre for Health Informatics Research & Development, Winchester www.chirad.org.uk (accessed March 2007)

[5] British Computer Society Health Informatics Forum www.bcshif.org (accessed March 2007)

[6] Special Quality Needs and Disability Act (SENDA 2001) www.opsi.gov.uk/acts/acts2001/20010010.htm (accessed March 2007)

[7] Mason A, Bishop P, et al. Developing a District IT Policy, Kings Fund/ Korner for Department of Health ,1983.

[8] Machan C, Ammenwerth E, Bodner T. Publication bias in Medical Informatics evaluation research : Is it an issue or not? MIE2006 Congress Proceedings, European Notes in Medical Informatics Vol II, No 2,2006.

[9] NHS Health Informatics Faculty www.informatics.nhs.uk (accessed March 2007)

[10] Roberts JM. The contribution made to the support of healthcare delivery and management by health informatics at a local and national level over a period of thirty years. Doctoral Thesis Teesside University,March 2006.

[11] BCS HI Forum, The Way Forward for NHS Health Informatics: Where should NHS Connecting for Health go from here? December 2006.

[12] UK National Health Informatics Collection www.uclan.ac.uk/nationalhealthinformaticscollection/ (accessed March 2007)

Address for correspondence

Beverley Ellis, HIU, Lancashire School of Health and Post-graduate Medicine, Harrington Building, University of Central Lancashire, PRESTON, PR1 2HE, UK bsellis@uclan.ac.uk

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

The EIPEN Project: Promoting Interprofessional Education in Health Professions

Joseph Liaskos, Antonis Frigas, Kostantinos Antypas, Dimitrios Zikos, Marianna Diomidous, John Mantas

Health Informatics Laboratory, Faculty of Nursing, University of Athens, Greece

Abstract

The Leonardo project under the name European Interprofessional Education Network (EIPEN) in health and social care, has been dealing with the challenges of Interprofessional Education (IPE). The EIPEN project tries to develop a transnational network of universities and employers in the six participating countries and at the same time to promote good practices in Interprofessional Learning and Teaching in health and social care. IPE provides opportunities for students and practitioners to learn with, from and about each other during qualifying and post-qualifying training and in their practice. IPE in health and social care includes the education and training of practitioners in human and animal medicine, dentistry, nursing, physiotherapy, occupational therapy, pharmacy and all other health professions including public and environmental health and health promotion, and social work. The outcomes of the EIPEN Project will provide means, material and guidelines for the enhancement of professional education in the multidisciplinary field of Health Informatics.

Keywords:

education, professional education, interprofessional education, multiprofessional learning

Introduction

According to the World Health Organization (WHO) Interprofessional Education (IPE) is identified [1] as an important component of primary health care. The Interpfofessional Education definition as given by CAIPE (Centre for the Advancement of Interprofessional Education) in 1997 [2] describes IPE as: “occasions when two or more professions learn from and about each other to improve collaboration and the quality of care”. The term multiprofessional education is also used to describe occasions when two or more professions learn side by side for whatever reason. The term interdisciplinary or multidisciplinary education or collaboration is referred to the combination and involvement of an assignment not necessarily working in an integrated or coordinated manner.

The interprofessional team in health has concentrated on two or at most three professions, primarily medicine, nursing and pharmacy. Interprofessional education has been invoked ever more frequently during the past thirty years to encourage collaboration in health and social care to help improve services, effect change and implement workforce strategies [3]. However, during the last decade, a renewed interest in IPE and an activity in health sciences programs appeared internationally. Inter-professionalism is very important while its importance is increasing; however there is not much literature on governance of Interprofessional Learning (IPL) [4].

Interprofessional education

Interprofessional education has been actively embraced in health professions since the 1950s, and in some cases earlier, and it is an evolutionary field of practice and research. It is seen as a way to develop collaborative practice among health and social care professions, and it plays a vital role in the patient-centred health services delivery approach. It contributes to collaborative practice and skills, aiming at patient safety. Teamwork is achieved through innovative methods of learning and practicing, improved ways of dealing with patient and technological complexity, enforcing professional relations and common understanding [3-7].

The implementation of interprofessional education is a difficult task for a various reasons: there are differences in prerequisites for admission to professional programs; the length of professional education; the extent and nature of the utilization of community and hospital resources for practice (clinical) education; students' freedom, or lack thereof, in the selection of professional courses; timetabling differences and conflicts across professional programs; faculty teaching loads; research interests of faculty; methods of administration within the various programs; and the powers vested in Deans of Faculties through statutory legislation, for example, through the power to appoint faculty members and to develop curricula [3,8].

Providing interprofessional learning experience that promotes teamwork and collaboration is a difficult task. There is a need to find academically acceptable mechanisms in order to measure the effectiveness of IPE activities. The existing attitudes of students, faculty and administration need to be changed in order to make IPE effective. The promotion of IPE as well as the measurement of its effectiveness requires that the students’ attitudes towards such work are assessed and evaluated.

Selected for best paper award.

Interprofessional education must confront particular challenges and needs seriously efforts in order to be successful. According to John H.V. Gilbert [8] these challenges include structural differences between faculty organizations; conflicting university and professional agendas; lack of adequate human resources to implement such programs, both within the university and across the community boundary; complex communication demands, within the university and with its community partners; rotation and replacement of team members; and lack of regular evaluation of interprofessional educational goals and programs.

IPE succeeds only when certain conditions are met: when the subject matter requires a team approach; when the effects of IPE can be clearly measured, for example, when critical reasoning skills are enhanced; when claims for resources to support IPE can be justified, that is, support for faculty and students is clearly necessary for success; when the skills being taught are within the competencies expected of a particular professional team; and when skills and knowledge can be explicitly taught and are clearly transferable, that is, those skills can be moved from one case to another [9].

Evaluation methods have to be developed in order to allow the assessment of the outcomes. These outcomes may include the patient, the process of interprofessional practice, individual professionals or agencies in which collaborations are carried out [10-12]. To exploit the benefits and outcomes of interprofessional education one must establish access to a wide range of resources of new knowledge and new skills. Many times, those benefits come through the shared respect and trust of the interprofessional partners who have been educated together in teams.

IPE and health informatics

Health Informatics is a multi-disciplinary field that deals with the collection, storage, retrieval, communication and optimal use of health related data, information and knowledge. The discipline utilises the methods and technologies of the information sciences for the purposes of problem solving and decision-making thus assuring quality healthcare in all basic and applied areas of biomedical sciences. Its domain covers computational and informational aspects of processes and structures in health care. Its aim is to study all the applications of informatics and computer science in Health Sciences (Medicine, Nursing, Dentistry, Biology, and Pharmacy) and health care. As a multi-disciplinary field it requires an educational approach aiming to collaborative learning. The different professionals who are graduates from specialised programs in health and medical informatics include physicians and nurses of different specialties, pharmacists, health care managers, computer scientists/ informaticians, engineers, e.tc. Moreover, practically all professionals in health care should, during their studies, be confronted with health and medical informatics education [13]. The development of educational curricula for Health Informatics follows an interprofessional approach.

In such an interprofessional educational environment (undergraduate or postgraduate), both teachers and students/ professionals from different disciplines have to be educated in a collaborative manner so as to develop the mutual understanding and respect among such multi-disciplinary groups, to enhance the existing and find new opportunities for shared learning and teaching, and advance the knowledge, skills, and attitudes of professional roles. For these aims, students will have to choose from a variety of learning activities such as team projects, which would provide opportunities for students to work together, tutorial courses focusing on a certain topic or area of health informatics theory or practice, exchange of experiences (e.g. observe team in action, discussion of real cases, placement of student or team in a real team of professionals in the clinical environment).

European Interprofessional Education Network (EIPEN) Project

The European Interprofessional Education Network in health and social care project, is supported and funded by EC Leonardo da Vinci Community action programme on vocational training (Project n^{o} 2005 UK/05/13/F/NT-162-335). It started on November 2005 and has two years duration until end of October 2007. The coordinator of the project is the Higher Education Academy Subject Centres, led by Health Sciences and Practice, based at King's College London (UK). In the network belong 16 partners from six countries. The European countries participating in the project are Finland, Greece, Hungary, Poland, Sweden, and United Kingdom, while the partners are universities, educational institutions, non governmental associations and centres related to interprofessional education [14-15]. From Greece two partners belong to the network: The Health Informatics Laboratory of the University of Athens and the, no budget holding partner, Greek Health Informatics Association (GHIA). The purpose of the project is to set up a network to develop and disseminate good practices in interprofessional education in health and social care in partner countries.

IPE provides opportunities for students and practitioners to learn with, from and about each other during qualifying and post-qualifying training and in their practice. IPE in health and social care includes the education and training of practitioners in human and animal medicine, dentistry, nursing, physiotherapy, occupational therapy, pharmacy and all other health professions including public and environmental health and health promotion, and social work. Interprofessional learning has been shown to lead to enhanced teamwork and health care. However, developing successful programmes and learning content and process is problematic. Thus partners desire to exchange teaching and management experiences as well as learning materials.

EIPEN has two interlinked aims: a) To develop a transnational, sustainable, inclusive network of people and organizations (universities and employers) in the six participating countries, and b)To share, develop, and promote effective interprofessional learning and teaching curricula,

methods and materials, good practices, for improving collaborative practice and multi agency working in health and social care. The model for network development is the successful UK Learning and Teaching Support Network, now the Higher Education Academy Subject Centres (www.heacademy.ac.uk), which uses a wide range of methods to disseminate and debate good practices in learning and teaching in Higher Education. The Subject Centres take account of the views of employers and other stakeholders, to sustain and extend their networks and to influence education policy. The network will be established over two years on a radial model. Each partner country will develop a website of resources and a programme of workshops and seminars that will link to a central interactive EIPEN web portal as well as regional and international events.

EIPEN goals

- Improved access to good practices, resources for learning and teaching, expertise, teaching materials and innovative case studies from universities, hospitals and other vocational training contexts in health and social care.
- A sustainable programme of events for presentations and debates concerning IPE.
- Dialogue with interprofessional education practitioners and peer consultation.
- New or improved sustainable national networks providing support for practitioners, teachers and students.
- Involvement of service users, students and policy makers.

EIPEN achieves these goals through the development of new networks within the health and social care sector through a programme of events at national, regional and transnational level:

- Interactive Web Portal with free membership registration, linked to national websites and resources
- Learning and teaching workshops and Seminars
- Transnational Steering Group Meetings
- International conference (2007 in Krakow)
- National learning and teaching resource databases linked to EIPEN portal
- Reports, presentations and publications
- Links with other International IPE groups

The products in two years includes the network, an interactive website with a resource database and a directory of people and organisation and discussion fora, and reports of national and international workshops and seminars on interprofessional training. The Transnational Network establishes and will be tested within 6 countries to ensure robust technology and systems and allow the design of procedures that will allow the resources of the network to be optimised. The aim of the project is to make the network available throughout the EU, but it is not possible to predict the specific longer term outcomes of EIPEN; these will emerge within the partnership. The web portal provides a facility for exploring other developments in health and social care practice and training, and the relationships being created increase opportunities for the exchange of ideas, staff and other partnerships.

Conclusion

IPE is innovative for professions, employers and education in health and social care. The innovation derives from the identification of opportunities for shared learning and teaching, the development of mutual understanding and respect among multi-professional and multi-disciplinary groups, the dissemination of interprofessional learning and teaching both in undergraduate and postgraduate education, and the advancement of knowledge, skills, and attitudes of professional roles. EIPEN develops new forms of networking within the health and social care sector for partners. EIPEN shares new products in the form of materials and methods for health and social care learning and teaching that will enhance the interprofessional education and learning in Europe.

Acknowledgments

EIPEN project is partially funded 2005-2007 by the European Commission as project No: 2005 UK/05/B/F/NT-162_335 in the Leonardo da Vinci Community action programme on vocational training. Project (network) partners are: Higher Education Academy, Health Sciences and Practice Network – King's College London (UK), Health Informatics Laboratory, Faculty of Nursing – National and Kapodistrian University of Athens (EL), Greek Health Informatics Association (EL), University and Polytechnic of Oulu (FI), Department of Mental Health Services – Oulu (FI), The Friends of the Young Association – Oulu (FI), Szechenyi Istvan University (HU), United Health and Social Institute – Gyor (HU), Department of Hygiene and Ecology – Jagiellonian University (PL), Occupational Medicine Centre of the Malopolska Region (PL), Karolinska Institute (SE), Centrum för Allmänmedicin Karolinska Intstitute (SE), The Primary Health Care Board – Stockholm County Council (SE), Centre for the Advancement of Interprofessional Education (UK), South West Peninsula Strategic Health Authority (UK), Steve Crane Consulting Ltd (UK).

References

[1] World Health Organisation (WHO). *Primary Health Care.* Report of the International Conference on Primary Health Care, Alma Ata. Geneva, 1978.

[2] Centre for the Advancement of Interprofessional Education (CAIPE). Interprofessional Education: What, How and When. CAIPE Bulletin No. 13. London, UK, 1997.

[3] Barr H. Interprofessional education: today, yesterday and tomorrow. UK: Learning and Teaching Support Network for Health Sciences & Practice, Centre for Health Sciences and Practice, 2003.

[4] Braithwaite J, Travaglia F. The ACT Health interprofessional learning and clinical education project: background discussion paper #1. The value, governance and context of inter-professional learning and practice. Canberra:, Braithwaite and Associates and the ACT Health Department, 2005.

[5] Lavin MA, Ruebling I, Banks R, Block L, Counte M, Furman G, et al. Interdisciplinary health professional education: a historical review. Advances in Health Sciences Education 2001;6(1):25-47.

[6] Atwal A, Caldwell K. Do all health and social care professionals interact equally: a study of interactions

in multidisciplinary teams in the United Kingdom. Scandinavian Journal of Caring Sciences, 2005;19: 268-273.
[7] Hall P. Interprofessional teamwork: Professional culture as barriers. Journal of Interprofessional Care. 2005; Supplement 1, 188-196.
[8] Gilbert JHV. Interprofessional education for collaborative, patient-centred practice. Nursing Leadership 2005;18(2): 32-38.
[9] Parsell G, Bligh J. The Development of a Questionnaire to Assess the Readiness of Healthcare Students for Interprofessional Learning (RIPLS). Medical Education 1999; 33: 95-100.
[10] Freeth D, Hammick M, Reeves S, Koppel I, Barr H. Effective Interprofessional Education: Development, Delivery and Evaluation. Blackwell Publishing, 2005.
[11] Freeth D, Hammick M, Koppel I, Reeves S, Barr H. Evaluating Interprofessional Education: a self help guide London: Higher Education Academy Learning and Teaching Support Network for Health Sciences and Practice, [Online] 2005 [cited 2006 Nov 25]. Available from: http://www.health.ltsn.ac.uk/publications/ occasionalpaper /occp5.pdf
[12] Freeth D, Hammick M, Koppel I, Reeves S, Barr H. A Critical Review of Evaluations of Interprofessional Education, Higher Education Academy Learning and Teaching Support Network for Health Sciences and Practice, [Online], 2003 [cited 2006 Nov 25]. Available from: URL: http://www.health.ltsn.ac.uk/publications/ occasionalpaper/ occasionalpaper02.pdf
[13] International Medical Informatics Association, Working Group 1: Health and Medical Informatics Education. Recommendations of the International Medical Informatics Association (IMIA) on education in health and medical informatics. Stud Health Technol Inform. 2004;109:226-43.
[14] European InterProfessional Education Network (EIPEN), [Online]. 2006 [cited 2007 Mar 28] Available from: URL: http://www.eipen.org.
[15] European InterProfessional Education Network (EIPEN), Centre for Health Sciences and Practice [Online]. 2006 [cited 2007 Mar 28]. Available from: URL: http:// www.health.heacademy.ac.uk/projects/eipen/

Address for correspondence

Joseph Liaskos, University of Athens, Faculty of Nursing, 123 Papadiamantopoulou str, Athens 115 27.
E-mail: iliaskos@nurs.uoa.gr

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Building ICT Capabilities for Clinical Work in a Sustainable Healthcare System: Approaches to Bridging the Higher Education Learning and Teaching Gap

Kathleen Gray, Jenny Sim

Biomedical Multimedia Unit, The University of Melbourne, Australia
Centre for Medical and Health Sciences Education, Monash University, Australia

Abstract

There is a recognised gap in information and communications technology (ICT) learning and teaching in higher education for entry-level healthcare professionals. This paper proposes a research model for understanding the dimensions of this gap. We describe methodological approaches to understanding present practices, identifying levers for change and learning by doing. We discuss issues faced in getting started and sustaining momentum on the research that is an essential prerequisite to effectively build the ICT capacity required by the clinical workforce in a sustainable healthcare system.

Keywords:

clinical informatics, competency-based education, education research, interdisciplinary communication, professional education, teaching.

Introduction

There is a recognised gap in information and communications technology (ICT) learning and teaching in higher education for entry-level healthcare professionals, and it needs to be bridged in order to build the capacity required by the clinical workforce in a sustainable healthcare system. Despite advances in educating specialist informaticians, in offering informatics electives to pre-qualification healthcare professionals, and in mapping details of ICT capabilities for healthcare professions, fundamental broadly based workforce capacity is not yet being established.

The contribution of ICT towards building a more sustainable healthcare system cannot be achieved by specialist informaticians alone, and relies upon active use of ICT capabilities in the professional practice of all clinicians. Clinical professionals' development of essential informatics capabilities is of local, national and international significance for evolving professional practice and education standards, and reform of healthcare system operations and management. Health informatics is an increasingly influential part of the working environment of "clinical staff including doctors, nurses, pathologists, pharmacists and other clinical professionals" [1]. The informatics capabilities of such staff are also a matter of interest to the public, as stakeholders in the healthcare system; increasing integration of ICT into healthcare has important implications for the quality of care, and for patient privacy and safety.

Improving the implementation of ICT in the healthcare system cannot be done only by providing these capabilities as optional extras for study by entry-level clinical practitioners; rather, it is necessary to embed informatics learning thoroughly in other aspects of professional learning and development [2, 3]. Examples of current educational interest in the learning and teaching of such capabilities in various clinical professions include allied health [4], dentistry [5], medicine [6], nursing [7] and public health [8]. But often such capabilities are addressed only in stand-alone ICT modules or elective units of study offered to students in their basic clinical training.

Nor can the healthcare system's need for technological transformation be met by learning and teaching capabilities simply profession-by-profession for clinicians. Calls to adopt a collaborative and multidisciplinary approach in promoting good health and well being [9] and to build ICT capacity across the entire health workforce [10] underscore the need for cross-professional learning and practice. ICT is recognised by a number of health educators to be an area of capability well suited to collaborative learning across professional boundaries [for example 11, 12].

Methods

A comprehensive research agenda is required to assess and address the higher education learning and teaching gap just described, so that we know what are the most promising approaches to pursue, in order to bring about the necessary educational change. This section proposes a three-dimensional model of the terrain to be bridged; it suggests a methodology for understanding the present situation, identifying levers for change and acting deliberately in response to this gap.

Present practices

One dimension of the gap is our lack of detailed knowledge about present practice in ICT education. It is important that those who educate future clinical professionals access and share research into "understanding where we are now … in order to clarify what practical steps are needed to move forward into the future" [13].

Clarifying the present situation with regard to ICT learning and teaching can provide important baseline data to allow prioritisation of research projects and evaluation of the effects of interventions. Key aspects of research to be done in this area include:

- Research into current roles / uses of ICT in educational settings for entry-level health professions: Across higher education, "There is enhanced use of educational technology … but all too often this emulates traditional didactic teaching and testing instead of promoting student curiosity and autonomy… In other words there is too often a poor alignment … between what is taught and the competencies students will need in their later lives and work settings" [14].
- Research into expectations about ICT as expressed in curriculum standards and professional accreditation processes: "Curriculum documentation can be read as giving in principle scope and support for the teaching and learning of essential clinical uses … However, documentation does not systematically address principles or processes for learning experiences that would scaffold or guide a student's transition …to those that are necessary or desirable in contemporary clinical practice" [15].
- Research into ICT applications in use in various clinical workplaces where students do their placement learning and hold their first jobs after graduation: A reality check is required to compare recommendations re informatics knowledge and skills that all clinical professionals *should* have [for example, 16] against the environments in which they actually work.

Levers for change

Another dimension of the gap is our lack of detailed knowledge about factors influencing ICT learning and teaching in higher education. We can locate obstacles to and levers for bridging the gap if we have a better understanding of the current dynamics of learning, teaching and educational provision:

- Understanding learning, especially understanding what learners need to be taught formally versus what they already know: For example "very little empirical research has actually questioned the Net Generation about their experiences with technology and worked with educational practitioners to determine the implications this has for Higher Education" [17].
- Understanding teaching, in particular the attitudes and experiences that determine orientations of key staff such as degree and year-level coordinators: For example, academics teaching core aspects of medical degree studies may find informatics "difficult to conceptualize" as a field of study and they may be equivocal about its inclusion in professional training [18]; and a nursing educator asks, "How can we expect faculty to transform nursing education for a type of practice that they have not experienced?" [19].
- Understanding what is involved in the provision of higher education, notably the pressures of globalisation, massification and privatisation on educational quality, and thus issues in competition with ICT education for attention in the operation of every clinical degree. In an example taken from physiotherapy, "The issue evoking most concern and comment is that of the ability of schools … and their professional clinical colleagues to continue to deliver appropriate clinical education within current resource constraints" [20].

Learning by doing

Another dimension of the gap is our lack of detailed knowledge about educational development prospects. Educational change is continuous and complex, and it is important that we recognise and capture the contribution of 'learning by doing' to what options there are for integrating ICT education into mainstream curriculum, and for finding interprofessional and inter-institutional ways of educating for ICT capabilities. This includes understanding what may be achieved by:

- Designing and implementing new learning and teaching resources and environments, such as plans "to integrate an entire enterprise-wide, electronic health record (EHR) system into the teaching curricula of nursing, physicians, pharmacists, and allied health schools, as well as health informatics and computer science" [21].
- Taking educational leadership roles in macro curriculum reform. Current health workforce shortages are producing many alternative scenarios for ways in which universities might partner with health services to provide frameworks for "experimentation and responsiveness in terms of preparing new types of health workers" [22].
- Strengthening research-led teaching or the teaching-research nexus, so that entry-level degree programs are appropriately informed by research into informatics – for example, there are a number of instances of research into using wireless handheld devices in clinical practice now being translated to enrich clinical teaching [23].

Such a research agenda needs to take a grounded theory approach. That is, it needs to iteratively and comparatively use data obtained from mixed methods of research - analysis of research literature and public documents, field observations and interviews, empirical studies; action research; individual and group evaluation by research participants - to provide an evidence base for the design and implementation of ICT learning and teaching for professional practice that will lead to systematic and sustained educational improvements. This approach is "uniquely suited to form the basis of research programmes that arise from theory grounded in the [health sciences] education experience, and then build toward implementation of practical educational innovations" [24]. Such an approach to research addresses recommendations for structuring clinical education [25], for developing authentic learning [26], and for supporting successful implementation of interprofessional education for collaborative practice [27]. This approach to research also accommodates the concept that educational development can overcome aspects of frag-

mentation and build a new professionalism in academics' working lives, through sponsoring a series of critical conversations [28].

Discussion

Preparing to advance upon a three-dimensional model using a toolkit of mixed methods to fill a major gap in clinical education knowledge and practice is a daunting prospect. It is no wonder the gap remains wide. In this section of the paper we review issues of when and where to start and how to sustain momentum, and we highlight some overarching considerations.

Of course, such work has been started, by many redoubtable educational researchers and practitioners, and has been in train for some time, however it continues to be hampered by issues of:

- The relatively low status of research into teaching and learning, especially in comparison with other research agendas in health and life sciences, and the consequently weak coordination and recognition of research in this field: this may be illustrated by the anomaly of strong emphasis on evidence-based practice in clinical work alongside its near absence in teaching and learning.
- The challenge of planning and resourcing complex ICT education research that is not just do-able, but meaningful in scope and scale, in order to make any significant difference to learning, teaching or curriculum: There are a myriad of political and logistical pros and cons to be weighed up in working on projects as a local initiative, versus on a state level or nationally, and / or in doing comparative studies nationally or internationally.
- The never-ending quest for sustainable long-term research partnerships with stakeholders external to the education system, such as ICT developers and vendors, healthcare agencies, health system funders: how is it possible to establish, maintain and evaluate these partnerships, and how to sustain this commitment alongside the full academic teaching and administration workload of many educational researchers?

Conclusion

Our view of the research to be done, the approach needed to bridge the knowledge and practice gap and the challenges faced in taking on this agenda rests squarely on eight key change lessons for higher education [29]:

1. "There are far more options for improvement or innovation than there is time or resources to address them."
2. "Change is not an event but is a complex and subject learning / unlearning process for all concerned."
3. "Enhancements in learning programs generate a need for improvement in the systems and infrastructure that underpin them."
4. "The most successful changes are the result of a team effort in which the most appropriate and best positioned people are involved in a process of action learning."
5. "The change process is cyclical, not linear."
6. "Change does not just happen – it must be led."
7. "Change is a mix of external forces and individual action."
8. "We must look outside as well as inside for viable change ideas and solutions."

This research needs to have a broadly-based view of the terrain, a robust methodological framework and mechanisms to share findings across boundaries and communities. Stakeholders must work together to keep this research agenda in the forefront in teaching and learning quality and research quality forums; in professional and industry forums on quality of care and workforce planning; in relevant non-governmental organisation and government forums. This work requires coordination and commitment among various stakeholders at multiple levels, and it requires substantial resources to ensure outcomes that are of significant benefit to the community, whether locally, nationally or internationally. Compounding the difficulty, Australian national classification systems for research fields do not enable ready identification of health informatics education research or give such research high visibility. However, it is only through seeking change in these conditions we will see real progress being made to bridge the higher education learning and teaching gap in building ICT capabilities for twenty-first century clinical work in a sustainable healthcare system.

The need for informatics education for non-specialists has been obvious for many years now [30, 31]. Our aim in this paper is to encourage the growth in strategic collaborations to take on all aspects of the associated research agenda over the next few years, in order to improve approaches to developing essential informatics capabilities in the entry-level education of clinical professionals. We welcome enquiries and approaches from others interested to work with us in this endeavour.

Acknowledgements

The authors wish to acknowledge influences on formulating the ideas in this paper that came from discussions during 2006 with Professor Evelyn Hovenga, Associate Professor Tangerine Holt and Dr Gregor Kennedy.

References

[1] Australian Health Information Council. Workforce capacity building national action plan. Canberra; 2003.

[2] Hersh W. The full spectrum of biomedical informatics research and education at OHSU. In: Haux R, Kulikowski C, editors. IMIA year book of medical informatics 2005: ubiquitous health care systems. Stuttgart: Schattauer; 2004. p. 167-72.

[3] Pearson A, Urquhart C. Health informatics education – working across professional boundaries. *Library Review.* 2002: 51(3/4): 200-10.

[4] Williamson D, Toth A, Grain H. Building health informatics capacity for the allied health professional workforce - case studies from the university sector. In: Proceedings of HIC 2005: Thirteenth National Health Informatics Conference;

HINZ 2005: Fourth Health Informatics Conference. Melbourne, Australia; 2005.
[5] Schleyer T. Dental informatics: A work in progress. *Advances in Dental Research.* 2003 17: 9-15.
[6] Blumenthal J, Mays B, Weinfeld M, Banks, M, Schaffer J. Defining and assessing medical informatics competencies. *Medical Reference Services Quarterly.* 2005 24(2): 95-102.
[7] Smedley A. The importance of informatics competencies in nursing: an Australian perspective. *Computers, Informatics, Nursing (CIN)*. 2005 23(2): 106-10.
[8] Karras B, O'Carroll P, Oberle M, Masuda D, Lober W, Robins L, Kim S, Schaad D, Scott C. Development and evaluation of public health informatics at University of Washington. *Journal of Public Health Management and Practice.* 2002 8(3): 37-43.
[9] Department of Education, Science and Technology (DEST). National research priorities: promoting and maintaining good health. 2005. [Monograph on the Internet] [Cited 2006 November 30]. Available from http://www.dest.gov.au/sectors/research_sector/policies_issues_reviews/key_issues/national_research_priorities/priority_goals/promoting_and_maintaining_good_health.htm
[10] Iliffe J. Building the IM & ICT capacity of the health workforce. In: Proceedings of the National Health Information Summit: transforming Australia's health system. 2005. p.121-3.
[11] Alpay L, Littleton K. Contexts for collaboration in healthcare education. *Health Informatics Journal*. 2001 7: 121-6.
[12] Carroll C, Rychlewski W, Teat M, Clawson D. Enhancing entrepreneurship and professionalism in medical informatics instruction: a collaborative training model. *Journal of the American Medical Informatics Association*. 2004 11(4): 317-9.
[13] Iliffe J. Building the IM & ICT capacity of the health workforce. In: Proceedings of the National Health Information Summit: transforming Australia's health system. 2005. p.121-3.
[14] Knapper C. Three decades of educational development. *International Journal for Academic Development*. 2003 8(1/2): 5-9.
[15] Gray K, Kennedy G, Judd T. Teaching health informatics to the net generation: a new baseline for building capability? Paper presented at MedNET 2006 Congress. Toronto, Canada. 2006.
[16] Garde S, Hovenga E. Australian health informatics educational framework. Australian College of Health Informatics. 2006.
[17] Kennedy G, Krause KL, Gray K, Judd T, Bennett S, Maton K, Dalgarno B, Bishop A. Questioning the net generation: a collaborative project in Australian higher education, In: ASCILITE 2006 Proceedings. Sydney: University of Sydney. 2006.
[18] Buckeridge D, Goel V. Medical informatics in an undergraduate curriculum: A qualitative study. *BMC Medical Informatics and Decision Making.* 2002 2: 6-16.
[19] Skiba D. Informatics competencies. *Nursing Education Perspectives.* 2004 25(6): 312.
[20] McMeeken J, Webb G, Krause KL, Grant R, Garnett R. Learning outcomes and curriculum development in physiotherapy. Australian Universities Teaching Committee / The University of Melbourne. 2005.
[21] Hovenga E, Louey-Gong D. Industry-government-academic partnership: the Australian story. In: Weaver C, Delaney C, Weber P, Carr R, editors: Nursing and informatics for the 21st century: an international look at trends, cases, and the future. Chicago. HIMSS. 2006. p. 206-9.
[22] Duckett S. Intervention to facilitate health workforce restructure. *Australian and New Zealand Health Policy.* 2005 2:14.
[23] Kho A, Henderson L, Dressler D, Kripalani, S. Use of handheld computers in medical education. *Journal of General Internal Medicine.* 2006 21(5): 531-7.
[24] Kennedy T, Lingard L. Making sense of grounded theory in medical education. *Medical Education.* 2006 40: 101-8.
[25] Henderson A, Winch S, Heel A. Partner, learn, progress: a conceptual model for continuous clinical education. *Nurse Education Today.*2006 26(2): 104-9.
[26] Barab S, Squire K, Dueber W. A co-evolutionary model for supporting the emergence of authenticity. *Educational Technology, Research and Development.* 2000 48(2): 37-62.
[27] Ginsburg L, Tregunno D. New approaches to interprofessional education and collaborative practice: lessons from the organizational change literature. *Journal of Interprofessional Care*. 2005 Supplement 1: 177-87.
[28] Rowland S. Overcoming fragmentation in professional life: the challenge for academic development. *Higher Education Quarterly.* 2002 56(1): 52-64.
[29] Scott G. Effective change management in higher education. *Educause Review.* 2003 38(6): 64-80.
[30] Staggers N, Gassert C, Skiba D. Health professionals' views of informatics education. *Journal of the American Medical Informatics Association*. 2000 7: 550-8.
[31] Coeira E. Medical informatics meets medical education: there's more to understanding information than technology. *Medical Journal of Australia.* 1998 168: 319-20.

Address for correspondence

Kathleen Gray
Biomedical Multimedia Unit
Faculty of Medicine, Dentistry and Health Sciences
The University of Melbourne VIC
Australia 3010
kgray@unimelb.edu.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

The Development of an Online Clinical Log for Advanced Practice Nursing Students: A Case Study

Jeffry S. Gordon [a], Ryan McNew [a], Patricia Trangenstein [a]

[a] *Vanderbilt University School of Nursing, Nashville, TN USA*

Abstract and objective

Three years ago at the Medinfo conference a prototype version of a clinical log for nursing students at Vanderbilt University was demonstrated. The purpose of the log is to document the types of clinical experiences the students are participating in as part of their academic program. We collected log data during that first year and received ongoing formative feedback from both students and faculty regarding its current feature set and desirable features for future implementations. Most of the requested new features have now been implemented. This paper describes some of the latest features of the clinical log, the advantages and disadvantages of ongoing development versus acquiring commercial products, and the procedures and results we have put in place to gather from faculty and students the features they want to see in the log. This paper also documents some of the data from early data mining.

Keywords:

education, nursing, graduate; nurse practitioners professional practice; students, nursing; computers, handheld/utilization; education, nursing, graduate/ methods; nursing education research; preceptorship

Introduction

Vanderbilt University's nursing curriculum is an Advanced Practice program designed to turn out Nurse Practitioners at the master's level in the following specialty areas: Acute Care, Adult, Family, Pediatric, Acute Care Pediatric, Midwifery, Psychiatric/Mental Health, Women's Health, and Neonatology.

Five years ago the PNP specialty faculty approached the informatics group about building an online clinical log to track their students' clinical experiences. They wanted a web-based version because it would provide more timely data to the faculty on the types of experiences their students were having. Until then, the students completed a notebook based paper log, sending it in twice a semester, and completed an op-scan questionnaire that was completed at the end of the semester. Op scan questionaires became increasingly popular in the early 1900s as an alternative to paper (Witzke et al., 1990) however, neither log type provided timely data. Notebooks could be lost in transit and it created a backlog several times a semester for faculty who had to read, make decisions on, and comment on the clinical experiences the students were having. McVeigh (1997) notes that paper logs are very inefficient. The faculty wanted the ability to check student experiences nightly to make sure there were no experience gaps. Gaps, if found, could then be dealt with in a timely manner (Alderson et al., 1999). They wanted the students to look for specific types of cases and log them as documentation that they had participated in them. The initial effort was rather primitive but it did provide the students with the ability to nightly log their clinical experiences and allowed their instructors to comment nightly on these experiences. Within two months, two other specialties (ANP, FNP) requested similar logs for them. This clearly escalated the development from "proof of concept" to the production of a professional quality web tool.

Currently the Vanderbilt clinical log supports students in 8 of the 9 specialties and will cover all 9 specialties next year. Initially we attempted to make all the specialty logs very similar. This would not only aid in coding the logs but it would make cross specialty comparisons a do-able procedure. Unfortunately, the significant differences in specialties and specialty requirements argued against keeping them very similar and as the years have progressed, even the basic features of the logs themselves have significantly diverged.

Methods

The initial logs were created by meeting with faculty in the specialty areas, demonstrating our existing log, and working from that. The datasets in the logs themselves were composed of generic data (time of patient meeting, type of insurance, gender of patient etc.), radio buttons indicating the level of participation in the encounter, checkboxes geared to the specific activities within the specialty, ICD9 code area, and a comment area. The student would treat each patient encounter as an individual record by entering the data above. The data could then later be queried by the student or the faculty member and summaries could be printed out. The faculty member could then write comments back to the student that would be included in the log. For all practical purposes, the only initial differences by specialty were the checkboxes identifying what procedures the student participated in.

After the first year, both faculty and students requested a large numbers of new features and additions they wanted

in the log. One area of immediate concern related to the checkboxes themselves. Standards and competencies from professional organizations were used to create the procedure checkboxes that was at the heart of the log. The log was created, to a great extent, to be faculty independent. It would be unacceptable to have a faculty member who was responsible for working on the log to leave, only to be replaced by another faculty member who wanted a completely different set competencies and experiences documented. The faculty examined the National Organization of Nurse Practitioner Faculties (NONPF) competencies in their specialty areas and designed their competency checkboxes from these competencies. Since NONPF was only creating some of those competencies during that time, it became apparent, particularly in the Psychiatric/Mental Health specialty that these checkboxes would have to be revisited after a year.

Results

We had three goals in the development of the clinical log.

1. The system had to be powerful enough to collect the right kinds of data. It had to meet the needs of the students and faculty and reflect the standards set by the professional organizations.
2. The system had to be easy to use. If it was difficult to navigate or cumbersome and time consuming to use the students would object. The faculty did not want the students to think of this as busy work.
3. The system had to be extremely reliable. Many of these students are not computer comfortable. They would never know whether the failure was at their end or our end but it would not matter. The frustration would be evident.

We met with each specialty of students three times formally during the year and with the faculty in each specialty at least one time formally per year. We made sure that students and faculty felt comfortable enough to approach any one of us at any time, through face to face meetings, email, telephone, or instant messenger.

The primary concern of the specialty faculty was to create a system that was easy to use by both students and faculty and where the students would not be bogged down nightly recording data as busy work. It is well documented in the literature that students are concerned about being overwhelmed with busy work while completing a log (Bardes et al, 2005) The faculty requested that a record should take no longer than 60 seconds to enter, and short of lengthy comments that goal has been met. The faculty also wanted a way in which students could flag certain records as important for their instructor to view. A checkbox was created that allows the student to indicate to the faculty member that this is a record of consequence. The students wanted to know when faculty commented on a specific encounter so an indicator was created on the student side showing which specific records had received faculty comment. The faculty members could even select the color of ink of the comment and tie that color to the flag indicator on the student side. While we assigned no importance to the color, and most faculty continued to write their comments in black, the additional color feature was a useful way for specific faculty to set priorities on comments.

The system had to have an easy to use login procedure. It was surprising that even well into the semester, students would forget their login names and passwords. In response, we created an email utility that would email them their lost password, but many students refused to use it. Since the university has a common username and password login for email, student records, and its course management system, we elected to switch the login procedures to that and this year not one student has contacted any of us about forgotten usernames and passwords.

System stability was a very critical feature. Many of these students and some faculty do not claim to be computer comfortable. Failures for any reason would cause them to balk at using the system and the students would then pressure the faculty to return to a paper log environment. Other than students forgetting their passwords, the system has never been down in three years. On several occasions faculty could not initially see a particular student but that was because the procedure we used to identify them by specialty incorrectly placed several students into the wrong specialty. Once those were identified (7 in three years) the problem was fixed manually. To summarize: The system had to be easy to use, powerful enough to provide students and faculty with the data they needed, and it had to be stable.

After about a year into development, several commercial products became available and we offered to consider switching to one of those applications... Faculty however, did not want to switch. The reasons they gave were: “we already know this log and it works.”, “why purchase something when we already have what we want”, “I like that I can make recommendations for changes and they will be implemented right away” (we quickly learned that rapid turn around on requests built strong loyalty for our efforts), “if our students have problems they know exactly where to get help”. Those were balanced against reasons for switching to a commercial product: We could get out of this business and focus our development resources elsewhere and there would be greater opportunities to share data across schools. As much as it was discussed, neither of these reasons was compelling enough to get the faculty to adopt a commercial product and the team realized that by customizing the log to their specialty requirements, we were building a product that they would use far into the future.

While great efforts were made to try to at least keep the logs of the various specialties parallel, it became apparent that this was not going to be possible. The Psychiatric/Mental Health specialty is an interesting case study on how difficult it became to keep the logs in synch. First, the PMHNP specialty uses DSM codes, not just ICD9 codes. Secondly, the manner in which they interact with their patients during their clinical hours is substantially different. Typically the other specialties’ students see any given patient one time. It is extremely rare that they would see a patient twice during their clinical experience. For these specialties, each encounter was equivalent to seeing a new patient. The PMHNP specialty’s students, however, see the

same patients over and over again week after week and they wanted the ability to "tie" the common records of a particular patient together. Retrofitting the log with this capability presented an interesting collection of challenges but they needed this for two reasons. It would make it easier to document patient progress over time if they could aggregate a specific patient's records in one place and they wanted a way to save their students' time by eliminating the need to check a host of demographic boxes on the same patient each time when the contents of those boxes essentially changed little from encounter to encounter.

Initially we provided a comment field for students to write their clinical comments on a particular encounter but the 5000 character limit was quickly discovered as inadequate for the PMHNP specialty. Furthermore the PMHNP faculty wanted their students to work off of a Microsoft Word created template that was not possible to code into the system. We then created an uploader that allowed the student to write their clinical comments in the Word template, save it under a specific record name that tied it to a particular encounter, and then upload the file to the server. The faculty side was then flagged to inform the instructor that a file attached to a specific encounter was available for viewing and comment. This year the PMHNP faculty wants the students to audio record the encounter. We have now created a methodology for recording the audio on their PDA, converting the file to mp3, and using the same uploader to attach the audio file to the appropriate encounter record.

Since certification requirements differ considerably between specialties, we had to create customized ways of tracking time and activities in an encounter. The Family PMHNP students for example, need to separate and track direct contact hours with patients under 18 versus direct contact hours with patients 18 and over, indirect hours related to time spent preparing for an encounter, and supervisory hours related to time spent conferring with a supervisor. Students were keeping this data on paper separate from the log. The instructors approached us with this issue and we immediately put it into the log. Other specialties, such and Midwifery, on the other hand, needed to track the number and types of births the student participated in.

Students approached us very early on and complained that they were in fact entering data on each patient twice: Once on paper after the patient had departed, and again online in the web interface that evening. Garrett (2006) identified the same issue. The students wanted a way to only enter data once and so we came up with a PDA solution to the problem. The student would enter the data immediately after the clinical encounter on their PDA, then they would synch the PDA nightly to avoid losing data. Once a week that synched data would be uploaded to the server and installed into the online database. We decided to force each specialty to use the web version only for one year, and then when the checkbox choices stabilized we would build them a PDA version. Constructing the PDA version would be difficult and we wanted to avoid making major changes in it each year. Currently we have PDA versions for 6 of the 9 specialties.

This year the PDA usage with the log has been as follows:

Table 1 – PDA Users

Specialty	Students	PDA users	Percent
ACNP	43	9	21%
ANP	35	6	17%
FNP	62	11	17%
PNP	34	19	56%
NMW	19	11	58%
PMHNP	28	0	0%

While 100% acceptance of the PDA solution was never expected, the relatively low percent of students using this technique was somewhat surprising. Garrett (Garrett et al., 2006) points out that the inherent restrictions on PDA can result in limited use. However, even with low numbers, the allegiance to this approach by the students practicing it was substantial. If a student started using the PDA environment to record their data, they stuck with it week after week.

Students gave a number of reasons why they chose not to use the PDA approach once they were given the option.

1. While PDA purchase was strongly encouraged and, in some cases required by each specialty, some students refused to make the purchase.
2. Some students balked at the $40 charge to purchase a PDA database program needed to make the PDA clinical log work. They opted to stay with the free web version.
3. Installing all of the software on the PDA to make it work was rather onerous. Some students claimed they couldn't get over that hump even though we created a web based Camtasia presentation that explained it in detail.
4. PDAs can be intimidating, particularly when they don't work the way the user thinks they should work and some students therefore just don't like PDAs.
5. Almost half of the students in the PMHNP program, for example, who commented on the PDA problems in their course evaluations, alluded to the lack of screen geography or difficulty in navigating the PDA clinical log application due to small screen space. This suggests that a tablet computer might overcome their reluctance here.

Last year none of the PMHNP students elected to use the PDA for data collection. When we asked them about that it was clear that the primary feature of the clinical log that they regarded as most important was the comment area and the corresponding ability to upload a Word file on the encounter to the record (Over 75% of the students were regularly uploading Word files with each log entry as

opposed to using the comment area in the log.). The PDA piece actually got in the way of their ability to do this. Comments of substantive length are cumbersome to write on a PDA and uploading a file works best when it is added to the record immediately after the record was created. Since PDA records were only uploaded once a week, the delay became an issue for these students. However, they do use the PDA to audio record their encounters. What this demonstrates is that people will use the technology they need to do their job as long as the technology helps them do their job, but they will not conform what they need to do to accommodate the technology if they perceive the technology is getting in the way. Since the PDA clinical log got in the way of the PMHNP students' ability to record the necessary data, they elected to just not use that capability.

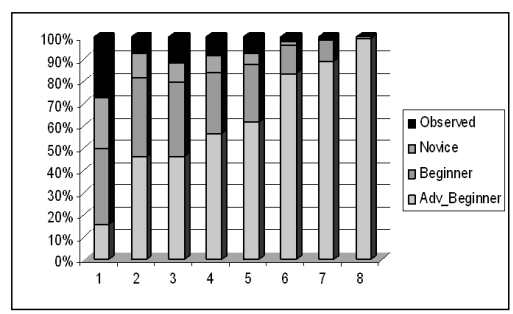
Perhaps the high percent of PDA users in the PNP and NMW groups is due to interested technologically literate and supportive faculty members who modeled behavior by using the PDAs for a variety of experiences in class. Since one of those faculty members has since left the institution, it will be interesting to see whether those numbers hold this year.

Early on in the development the specialty faculty members were asked what types of data queries they would want for both themselves and their students. It was necessary that they be very careful what they asked for because these queries are cumbersome to write. Because of limited resources, it was important that they not request queries they ultimately would not use... Rather than develop a series of queries the team created a tool to output the data into a spreadsheet table (Excel format) and taught the faculty and students how to analyze the data from there. The faculty appreciated the opportunity to learn Excel since they felt it could assist them in their research. They also said that since their students are becoming advanced practice nurses who will participate in research they too should learn Excel as a research tool. However, even with hands-on experience learning Excel in the computer lab, some students were just having a lot of trouble picking it up after a single training session. In response, the team created a web based animated training module using Camtasia and made that available on demand. Students could then go through the narrated training module whenever they were working on their analyses. In two of the specialties, the faculty showed the Camtasia presentation in class. The students commented that with the Camtasia created presentation, preparing the final analysis was very easy. The NMW specialty faculty, however; decided they wanted to create special queries outside of Excel. One of the students in the NMW program had an undergraduate degree in Computer Science and offered to program the queries. She was granted access to the DSN name, database tables, and variable names. She said she appreciated the chance to renew her skills in the area. Within several months she had a variety of queries programmed directly into the environment. This year, the NMW faculty wanted a significant number of changes to their log and we were concerned that the queries the student developed would no longer work. However, her code was well documented and it was easy to make the necessary changes to it. The other specialties have now seen the NMW queries but, much to our surprise, have not asked for them, even though we have offered building them. Apparently they still prefer the Excel approach and appreciate that we are teaching this lifelong skill to their students.

Students report that by using the output from the clinical log as a portfolio, they appeared to have an advantage in searching for the positions they wanted. One psychiatric student was able to demonstrate the types of patients he had worked with to such a degree that they offered him the position immediately. Others reported similar encounters. Several students asked if we could keep their accounts in the log active because they wanted to keep an ongoing portfolio of their more interesting cases and wanted documentation to support their activity.

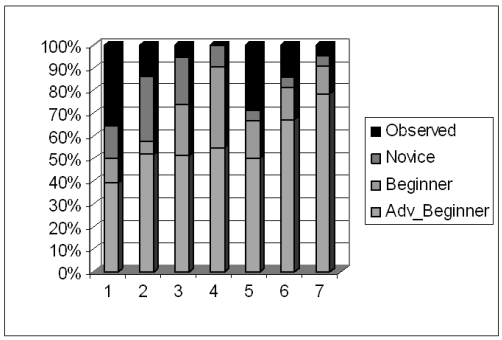
Faculty are now beginning to discover how to data mine the data tables and are seriously looking at the kinds of experiences their students are having across years. The overall data shows that the students definitely do gain clinical experience and confidence through the year. In the log the students evaluate their involvement in each encounter on a 4 point scale: (Observed, Novice, Beginner, up to Advanced Beginner with descriptor adjectives to enhance reliability of assessment). Figure 1, below, from the Family Nurse Practitioner (FNP) specialty, clearly shows that the students move, almost linearly from mostly observing to becoming more advanced in their participation over their 8 month clinical experience.

Figure 1: Type of involvement per encounter by month (FNP)



The relatively non-linear shape of the curve, below, in the Acute Care Pediatric Nurse Practitioner (PNP-AC) specialty is now serving as the basis for discussion on the types of clinical experiences this specialty is having, even though the general trend is toward more advanced experiences. Why, for example, is it relatively flat from months 2 to 5?

Figure 2: type of involvement per encounter by month (PNPAC)



We are now going to embark on a serious look across years and across specialties to guarantee that we are meeting the needs of the disciplines. We currently have approximately 490,000 individual patient encounter records in the system spanning the three years. As we add our final specialty and offer to make the log available to our partner that teaches the anesthesiology specialty, we anticipate adding over 200,000 records per year to the system.

Our next steps in the development process consist of altering the procedures we use for the PDA side, exploring alternative input devices, such as the intelligent pen and java enabled cell phones, and stabilizing the type of data collected to better allow for year to year comparisons. At the request of the program faculty, the PMHNP PDA log will be abandoned. Instead they want us to facilitate using the PDA to better audio record the patient encounters their students have and we will spend our development time with them promoting that. The rest of the specialties want the PDA piece streamlined with the data available for faculty review nightly instead of once per week. We are currently looking at two different technologies to accomplish that task. Finally, there is the ongoing conflict between standardizing the data to facilitate program to program evaluations and the customizing each specialty area to more exactly meet the needs of each particular program. This next year there will be an attempt to create more standardization where possible, but after that we are going to let the data fields stabilize for about two years to permit year to year comparisons and analysis.

Conclusion

The development of the clinical log has been an exceptional team building experience between faculty in the specialty areas and the informatics faculty and staff. Creating a professional quality "home-grown" log has given the school an excellent example of what the skills of the informatics area can accomplish. The faculty has been very pleased with the quick response they get when they want a new or improved feature and now rely on the informatics team for all of its technological needs. In return, we has learned a great deal about what works and what doesn't with students by specialty area and has developed a close professional relationship with the specialty faculty. We have been surprised at how rapidly the faculty has adopted the log. Based on prior work we anticipated relatively long lead times to change the culture (Nierenberg et al., 2007) but that turned out not to be the case. The faculty is now beginning to data mine the log to improve the clinical experiences of their students and perhaps alter their clinical content as well as modify the focus in their didactic teaching. Program alteration does not come automatically, however, and effort must be made to use the logs to improve the structure of student learning (Dolmans et al, 1999) This experience has translated into many other additional projects. We are now using the log technology for a grant supported cell phone based messaging system to encourage smoking cessation. We have created many online assessments, surveys, and questionnaires for the school and associated organizations. Finally, we are exploring new technologies, (such as intelligent pen based data collection systems and java enabled cellular phones) to accomplish many of these tasks.

References

[1] Alderson, T.J., Oswald, N.T. (1999). Clinical experience of medical students in primary care: use of an electronic log in monitoring experience and in guiding education in the Cambridge community based clinical course. Medical Education, *33*, 429-33.

[2] Bardes, C.L., Wenderoth, S., Lemise, R., Ortanez, P., Storey-Johnson, C. (2005). Specifying students-patient encounters, web-based case logs, and meeting standards of the liaison committee on medical education. Academic Medicine, *80(12)*, 1127-32.

[3] Dolmans, D., Schmidt, A., Beck, J., Beintema, M., Gerver, W.J. (1999). Does a student log provide a means to better structure clinical education? Medical Education, *33*, 89-94.

[4] Garrett, BM, Jackson, C. (2006) A mobile clinical e-portfolio for nursing and medical students, using wireless personal digital assistants (PDAs). Nurse Education Today, *26(8)*, 647-654.

[5] McVeigh, S., Williams, R., Pangaro, L. (1997). An internet home page to log students' contacts. Academic Medicine, *72(5)*, 417-18.

[6] Nierenberg, D.W., Eliassen, S.M., McAllister, S.B., Reid, B.P., Pipas, C.F., Young, W.W., Ogrinc, G.S. (2007). A web-based system for students to document their experiences within six core competency domains during all clinical clerkships. Academic Medicine, *82(1)*, 51-73.

[7] Witzke, D.B., Koff, N.A., McGeagh, A.M., Skinner, P.D. (1990). Developing a computer-based system to document and report students' clinical patient encounters. Academic Medicine, *65(7)*, 440-41.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Personalized Case Driven Parental Education Informatics in the NICU

John Chuo[1], Pavel Sherman[2], Claire Drain[1], Casimir Kulikowski [2]

[1] *Department of Pediatrics, Robert Wood Johnson Medical School - UMDNJ, Bristol Myers Squibb Children's Hospital at Robert Wood Johnson University Hospital*
[2]*Department of Computer Science, Rutgers University, USA*

Abstract

Neonatal Intensive Care Units (NICUs) are foreign and intimidating to parents of premature infants. The high levels of anxiety and stress they can produce needs to be reduced by thoughtful advice from healthcare providers (HCPs), to educate parents about their child's condition. Unfortunately time constraints often limit HCPs to only a few minutes with each baby's parents daily – only enough to convey critical information at a high level and with limited depth. Parents searching for more detailed information themselves in the literature over the web have sometimes reported disappointing experiences. We are proposing to improve parental education by patient-centric web methods leveraging the electronic patient record with internet and cell phone technologies. This can be an important informatics resource, complementing and enhancing face-to-face communication through personalization of education and advice to the parents.

Keywords:

Neonatal Intensive Care Unit (NICU), Personalized Parent-Patient (PPP) model, personalized parental education, electronic health record (EHR), Health Care Providers (HCPs)

Introduction: NICUs and Parents

The NICU experience is an emotional and confusing time for parents; what many families expected to be a time of joy and happiness has turned into an event filled with uncertainties. Many parents entering the NICU feel overwhelmed by the intensive care unit environment, including the technology, language used, and the pace of events [1,2]. This, coupled with their own sense of loss of expectations for the birth and future of their child, can lead to miscommunication between staff and families. What may seem "normal" to the NICU staff, can feel anything but normal to the family. An important way to make the experience most constructive is to develop a trust relationship between families and their HCP through effective communication [3,4,5].

The informational asymmetry that exists between the medical staff and their patient families is a source of anxiety [5]. While every individual has their own communication and information processing styles, what is important to parents is that they have access to all of the information that they want to have, when they want it, and that they be allowed to absorb the information in their own time. Withholding information, even under the pretense of sparing the family pain, or perhaps unnecessary anxiety, does not create a trusting relationship between the staff and the family [2]. The Bayer Institute for Health Care Communication identifies education along with empathy, engagement, and enlistment as one of the key factors of successful communication. Learning and understanding is best done in steady short segments, rather than episodic large segments [9]. Benefits of an educated parent are less anxiety, less depression, shortened NICU stay, and increased satisfaction with the medical staff care [3-7].

The high acuity nature of the NICU prevents the HCP from spending the necessary time needed to provide optimal patient education. In one study, parents feel that important information was not relayed satisfactorily almost 50% of the time (i.e. either too much, too little, or not explained at all) [10]. We propose a solution for optimizing parental education in the NICU by leveraging technology to educate parents 'on the fly', engaging the HCP only for critical tasks such as content verification before sending to the parents. The system should map existing information in the EHR to existing peer reviewed knowledge resources. Our system (1) matches personal content in the EHR to existing knowledge resources (KRs), (2) presents those KRs to HCPs for approval before sending them to parents via a secure web portal, and (3) uses cell phone technology to alert HCPs and parents of new information available for sending and viewing, respectively. This strategy aims to enhance the personal meetings between HCP and parents by (a) preparing parents before the meetings with background information, (b) giving additional information that may clarify questions that parents may think of after the meetings but are unable to ask HCPs immediately, (c) expediting communication between the HCP and parents regarding important and unexpected issues that may arise during the baby's clinical course.

Table 1 Neonatal Clinical Scenario

Event		Action by HCP using EMR in the NICU			EMR Test results
Date Time	(#) Description	Add to patient's problem list	Orders	Notes	
Day 1 8 AM	(1) Infant delivered by normal spontaneous vaginal delivery with 1, 5 and 10 minute apgar scores of 9/9/9. Prenatal labs significant for GBS+,Hepatitis B unknown. Baby with respiratory distress.	Prematurity, Rule out sepsis, Hepatitis status unknown, Surfactant Deficiency Syndrome	*Test Order:* hemogram, blood culture, c-reactive protein, chest x-ray *Med Order:* Ampicillin, Gentamycin,Hepatitis Immunoglobin / vaccine	CPAP	
9 AM	(2) CXR c/w RDS,give 40% O_{2}		*Med Order:* Surfactant	Intubation*	
1 PM	(3) Hypotension, wide pulse pressures,		*Test order:* Echocardiogram		
3 PM	(4) Hypotension		*Med order:* Indomethacin		*Echo* - PDA
Day 2 8 AM	(5) Normotensive,Infant extubated and placed on CPAP			Extubation to CPAP	
1 PM	(6) Started feeds- breastmilk		*Feeding order:* start BM 1 ml every 3 hours via NGT		
Day 3-20	(7) Increased Apnea episodes Jaundice	Apnea Jaundice	*Med order:* Caffeine, Phototherapy		
Day 21 8 AM	(8) Routine hemogram significant for anemia	Anemia			Hematocrit = 25
Day 49 8 AM	(9) Ready for discharge		*Order:* Discharge		

Methods: Clinical scenario

Setting. Currently, physicians in our neonatal intensive care unit at the Bristol Myers Squibb Children's hospital enter patient information into an electronic health record system named NeoData, by Metasoft Systems, Inc. The Eclipsys 3000 TDS system is used to enter orders and view test results. Patient problem lists, orders, and test results will be obtainable by our software system. We will illustrate how our system can help educate patients' parents in our NICU. Table 1 highlights events for a typical neonate over a timeline spanning three common phases of a patient's hospital stay: Admission, Interim, Discharge.

Synopsis. TJ is a premature baby boy born to a 39 year old mother at 29 weeks gestation. The apgar scores are 9 at 1, 5 and 10 minutes after delivery. His mother's prenatal laboratory result is positive for Group B Streptococcus and her Hepatitis status is unknown. TJ's immediate medical issues are Surfactant Deficiency Syndrome (also known as Respiratory Distress Syndrome) and Patent Ductus Arteriosus, resulting in respiratory distress and hypotension. An immature immune system makes TJ vulnerable to infections;jaundice and apnea also occur. His stay in the NICU also involves extensive routine care issues that support proper growth and maturation until discharge.

Admission period. TJ's admission period (day 1-2) is one of the most clinically telling, data intensive and anxious moments. Perhaps the most important and longest HCP-parental meetings occur at this time, typically lasting 30 minutes. During this time, TJ's HCP team must not only orient his parents to the NICU environment, but also discuss prematurity and its implications, its many acute events, and be sensitive to the parent's level of confusion, stress, and medical unawareness – in TJ's case, the numerous topics are respiratory distress, sepsis, antibiotics, various diagnostic lab tests, and Hepatitis prophylaxis (event 1). The nurse at the meeting must review general nursing procedures and routine regulations as well as the operational rules of the NICU. Not surprisingly, when parents are later asked about what was discussed, many can not recall key points of the content [8]. Reasons for difficulties in knowledge retention are (1) too much, too soon, too complicated (2) emotional interference, and (3) lack of HCP follow-up.

A detailed discussion of all of TJ's acute clinical issues (prematurity, sepsis, respiratory distress syndrome) would be overwhelming. An HCP-given high level overview in lay terms of current clinical issues and forthcoming expectations is effective for the short term, but needs to be followed up by detailed information and opportunities for parents to ask follow-up questions after 'digesting' the news – all done using a family centered approach [13]. Our system enables ongoing delivery of important information, complementing what is learned during these meetings, especially the initial one, so that they can be better prepared to ask important questions in subsequent meetings.

Interim period. The interim period can be both long and unpredictable, testing the emotional fortitude and patience of parents and HCPs. In most cases, meetings with the HCPs are sporadic and dictated by clinical events rather than regularly scheduled updates. In one study, only 26% of the surveyed parents report talking to the neonatologist on a daily basis [10]. Without meaningful and well timed dialogues with HCPs, parents can easily be misinformed by the internet and non peer-reviewed literature. A study by Dhillon et al reveal that only 10% of parents feel medical information found on the internet is reliable and up to 24% found the information to be over-frightening [12]. Shortly after birth, TJ's respiratory distress worsened and prompted emergency intubation and placement on a ventilator. Soon he became hypotensive (events 3-4) and an echocardiogram revealed a patent ductus arteriosus, which

is a common life threatening condition causing hypotension. The treatment involves risky medication therapies like Indomethacin. For events 2-4, although physicians would prefer updating parents directly, having more detailed information available for the parents to review after being updated is important in the education process. Complex conditions such as apnea, anemia and jaundice could be signs of either life threatening problems such as intraventicular hemorrhages and biliary atresia, but more often are part of the normal stages of development that require patience and time. Preparing parents to expect these non-emergency, anticipated conditions ahead of time will reduce anxiety and build trust.. Educated parents tend to feel less vulnerable, more empowered and engaged [14].

The overall background of stress during this period is riddled with episodic roller-coaster moments of relief and worry that is dictated by the clinical picture. Most communicated information during this period falls into one of three categories: (1) diagnosis and prognosis (events 1,7,8), (2) policies and procedures (events 2,5), and (3) routine care (events 6,7). The timing of information delivery is important to patient treatment and parental care. Early awareness of important test results such as the presence of a hemodynamically significant Patent Ductus Arteriosus can expedite communication between HCP and parents so that time sensitive treatments can begin sooner (event 2). Our software would enable this function by sending a text message to the HCP as soon as the result appears in the EHR.

Discharge. The discharge period is one of the happiest yet intimidating times, marking the beginning of a new phase in TJ's care in the hands of his parents. Proper parental education is more important now than ever to TJ's well-being. Unfortunately, the discharge meeting has similar time and content constraints as the initial one that occurred nearly 2 months earlier. Researchers show that parents who have been educated all along about their infant's clinical issues, as our system would permit, are more prepared to take over as the baby's primary caregiver [15,16]. Following the same delivery strategy, the HCP will typically give TJ's parents an overview – this time of his special needs at home. The HCP will introduce issues that were not relevant to his hospital stay but important in the 'outside world' such as TJ's vulnerability to respiratory synctius virus infections and associated complications - therefore the need for medication prophylaxis. Proper parental education and understanding of TJ's developmental needs (i.e. sensory stimulation, engagement) at home would maximize his potential for a normal life – the discharge meeting only begins the education. Our system will supplement the issues discussed with additional information that the parents can review, so they can better learn from their pediatrician and TJ's development care team.

Proposed system design and functionalities:

Figure 1 illustrates the major components of our proposed system. A knowledge repository consists of reliable information to be delivered to parents/guardians and healthcare providers; a personalized parent-patient (PPP) model matching neonate events from the patients' electronic health records to parental roles and their information needs; an algorithm that personalizes the information delivered to parents/guardians; and a notification component for informing the appropriate parties when new information becomes available or previously delivered information is updated.

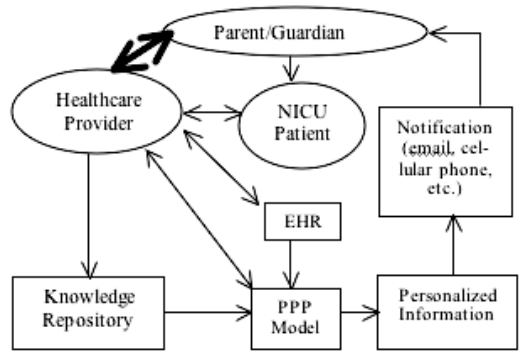


Figure 1: System Design

The knowledge repository (KR) stores all documents to be delivered and are approved by a staff member, which are reviewed and reapproved subsequently. Entering a document to the KR restarts the personalization algorithm to compute a relevance score of a new document for each PPP model. The document is marked for delivery for each PPP model for which the document's relevance score exceeds the threshold. A healthcare provider is then notified that new information is available for delivery, and on approval, the parent/guardian is notified that they can access it. The PPP model takes EHR data and maps it through its meta-data schema to categories of information delivery. Updates to an EHR will trigger the personalization algorithm to re-compute the relevance score for each document in the KR. After this is completed, the articles are recommended, approved, and delivered to the parents/guardians and/or healthcare providers associated with the appropriate child's record.

The relevance score for each document is based on weights developed by neonatologists and NICU staff based on how each document is triggered by the context of a neonate event, and answers a typical patient-parent information problem. Personalization comes from model matching the metadata of specific relevant documents to the metadata from the PPP models. If both a document and a PPP model contain the same term in their metadata, then the relevance score of the document increases by its weight (representing the informative urgency of the information for the PP context). If the document contains too many keywords not found in the PPP model, then the document's relevance score decreases by an "unlikely information" weight. Otherwise the score is left unchanged. After each document has been assigned a relevance score for each PPP model, those documents with relevance scores above a certain threshold will be queued

for approval for delivery to parents/guardians. The initial prototype is implemented with this procedure. We are currently experimenting with more sophisticated personalization algorithms to help improve the accuracy of the relevance score for a set of representative neonate scenarios (including TJ's), while reducing the burden on the HCP reviewers and approvers by summarizing displays and clusters of indices to the (as yet few, brief, and to-the-point) references. This raises many issues of informatics research we briefly address in the concluding section.

Table 2 shows how the system can improve communication between the parent and TJ's HCP. As the details of each event are recorded in the EHR, the metadata associated with each entry is compared with the metadata of the knowledge resources according to the personalization algorithm. The knowledge resources matching the EHR are queued for delivery to TJ's parents. The HCP may also choose to manually deliver certain documents from the KR. Messages are sent by cell phone to the parents immediately after the documents are approved. In some cases, the physician may not be aware of relevant updates to the EHR and the system will alert the physician of this, increasing communication efficiency. We now give an example of how the system can be used to deliver information to TJ's parents.

Table 2: Knowledge Resources for Clinical Scenario

	Knowledge Resource		**SMS Messaging**	
Event (#)	**Automatic (EHR guided)**	**Manual**	**To MD**	**To Parent**
1	NICU orientation, Prematurity, sepsis, Hepatitis prophylaxis, RDS, CPAP, CBC, CRP	Apgars, Group B Streptococcus, Blood transfusions	X	X
2	Intubation, Surfactant			X
3	Echocardiogram			X
4	PDA, Indomethacin		X	X
5	Extubation, CPAP			X
6	Store breastmilk, NGT			
7	Jaundice, Phototherapy, apnea, Caffeine	Kangaroo care		X
8	Transfusion			
9	Discharge instructions			

Discussion and conclusion:

We outline an informatics system that augments personal communications between parents and health care providers (HCPs), so that the information they "pull" from many sources can be channeled and interpreted in the context of their child's problems. The personalized technology of the "push" system will allow pertinent information to reach parents, and address their child's clinical status, in the most time sensitive manner. As we discussed, early and continuous education of parents has proven to be very effective in a NICU setting. In TJ's case, early presentation to the parents of a commonly anticipated clinical scenario during the NICU orientation, such as the likelihood of blood transfusion, will address several issues. It will allow TJ's parents time to assimilate and ask about the need for such a procedure, raising issues about TJ's safety and comfort, and the safety of the blood supply [17]. By presenting all of the information to TJ's parents as early in the process as possible, it will also allow the family time to find donors, and have them screened, for a directed blood donation, if that is something they want. Blood transfusions require parental consent. Having addressed all such relevant concerns earlier during TJ's course of treatment will allow the HCPs to care for TJ, and his family, in the most efficient way possible.

In a situation of greater emergency, such as TJ's need for intubation, the "push" system is also an effective tool to alert parents to the fact that their child's clinical status has changed. While we would not advise alerting parents of intubation via the system, it does provide the HCP with the opportunity to reach the parents and let them know that they need to call the unit to speak with their attending physician as quickly as possible. The system will also allow parents to be pushed information regarding the reasons for TJ's intubation, and the potential side effects of long term oxygen use. While we would not expect parents rushing to their child's bedside in an emergency situation to stop and read the information provided, it is reasonable to assume that they would have questions regarding the above issues. The "push" system will allow the HCP to anticipate common concerns of parents in this situation, and to address those concerns as quickly as possible. By sending the information to the parents electronically, the parents will be able to read and digest the information, and its implications, when they are ready. In this format the parents will also be able to revisit the information as often as needed, until they are comfortable with their own understanding of TJ's situation.

Currently, HCPs and administrators are the main users of EHRs. This proposal suggests leveraging the EHR as part of a personalized educational tool that uses physicians to direct the pushing of relevant educational knowledge to NICU parents. The implementation challenges involve (1) validation of knowledge resources, (2) appropriate mapping of KR to EHR content via the PPP model, (3) HIPPA and security, (4) integration into the already overtaxed physician workflow, (5) parental literacy, and (6) implement integration with EHR and other legacy systems. Our system will integrate into everyday workflow with mini-

mal perturbation of existing workflow by sending information to patients "on the fly" as they are entering orders, notes, or updating problem lists.

There are many informatics challenges in developing the system we have outlined here for the NICU. The personalization of information over multimedia databases can involve multiple steps of extracting reference literature, summarizing it, and matching it to a query based on contexts of extraction mapped to clinical databases, as investigated and carried out in the PERSIVAL project over a considerable period of time [18]. Many issues of text mining, semantic information modeling and query definition and refinement are involved in ways that present considerable challenges for a full-fledged system design. We have outlined a scenario in the NICU which, with others, are serving our group as use cases in the design of a prototype system for NICU-specific customization of information [19] in a patient-centered manner [20], but that will address issues unique to this health care context, where parents and guardians are the ones being communicated with, and where often inter-cultural and language issues must also be addressed [21]. The present paper summarizes some of the NICU specific factors that make this project such an important health informatics challenge to address.

References

[1] Young SR, Watson MA, Corff KE, Odle P, Hasse J, Bowerman JL. Parent stress and coping in NICU and PICU. J. Pediatr Nurs. 12(3): 169-77; 1997

[2] Conner JM, Nelson EC. Neonatal intensive care: satisfaction measured from a parent's perspective. Pediatrics 1999; 103(1): 336-349.

[3] Davis DL, Stein MT. Parenting Your Premature Baby. Golden, CO: Fulcrum, 2004.

[4] Mann MY, Lloyd-Puryear MA, Linzer D. Enhancing communication in the 21st century. Pediatrics 2006; 117(5): S315-S319.

[5] Bass LS. What do parents need when their infant is a patient in the NICU? Neonatal Network 1991; 10(4): 25–33.

[6] Worchel FF, Prevatt BC, Miner J, Allen M, Wagner L, Nation P. Pediatrician's communication style: relationship to parent's perceptions and behaviors. Journal of Pediatric Psychology 1995; 20(5): 633-644.

[7] Melnyk BM, Feinstein NF, Alpert-Gillis L, Fairbanks E, Crean HF, Sinkin RA, Stone PW, Small L, Tu X, Gross SJ. Reducing premature infants' length of stay and improving parents' mental health outcomes with creating opportunities for parent empowerment (cope) neonatal intensive care unit program: a randomized, controlled trial. Pediatrics 2006; 118(5): e1414-e1427.

[8] Wereszczak J, Miles MS, Holditch-Davis D. Maternal recall of the neonatal intensive care unit. Neonatal Network 1997; 16(4): 33–40.

[9] Barbara Klug Redman, PhD, RN, FAAN . The Practice of Patient Education, 10th Edition - A Case Study Approach. Elsevier Publishing, 2006

[10] Kowalsji WJ, Leef KH, Mackley A, Spear ML, Paul DA. Communicating with parents of premature infants: who is the informant? J of Perinatology (2006) 26, 44-48

[11] Koh THHG, Jarvis C. Promoting effective communication in neonatal intensive care units by audiotaping doctor-parents conversations. International Journal of Clinical Practice 1998;52:27-9.

[12] Dhillon AS, Albersheim SG, Alsaad S, Pargass NS, Zupancic JAF. Internet use and perceptions of information reliability by parents in a neonatal intensive care unit. J Perinatol 2003; 23: 420–424.

[13] Harrison H. The principles of family-centered neonatal care. Pediatrics 1993;92:643-50.

[14] Cullen R. Empowering patients through health information literacy training. Proceedings from World Library and Information Congress: 70th IFLA General Conference and Council. Aug 2004.

[15] Splingaire JM. Readiness for Hospital Discharge: Prdictors and Outcomes in NICU parents. http://stti.confex.com/stti/bcscience38/techprogram/paper_26745.htm

[16] Hospital Discharge of the High-Risk Neonate- Proposed Guidelines. Committee on Fetus and Newborn. American Academy of Pediatrics

[17] Ringer SA, Richardson DK, Sacher RA, Keszler M, Churchill WH. Variations in Transfusion Practice in Neonatal Intensive care. Pediatrics 101:194-200; 1998

[18] McKeown KR Chang SF Ciino J Fenier sk Friedman C Gravano L Hatzivassiloglou V Johnson S Jordan DA Klavans JL Kushniruk A Patel V Teufel S PERSIVAL, A System for Personalized Search and Summarization over Multimedia Healthcare Inforamtion. Proc. JCDL '01: 331-340, June 2001.

[19] DiMarco C Hirst G Wanner ? Wilkinson J HealthDoc: Customizing patient information and health education by medical condition and personal characteristics. 1st Int Wkshp on Artificial Intelligence in Patient Education, Glasgow, 1995.

[20] Gustaffson DH Hawkins R Boberg E Pingree S Serline RE Graziano F Chan CL Imapact of a patient-centered, computer-based health information/support system. Am J Prev Med 1999; 16: 1:9

[21] Piette JD Patient Education via Automated Calls 1999 J Prev Med 1999; 17:138-141.

Address for correspondence:

John Chuo, MD, MS
48 Plantation Rd. Whitehouse Station, NJ 08889, USA
chuojo@umdnj.ed

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Conceptual Model of Health Information Ethics as a Basis for Computer-based Instructions for Electronic Patient Record Systems

Mihoko Okada[a], Kazuko Yamamoto[b], Kayo Watanabe[c]

[a] *Department of Health Informatics, Kawasaki University of Medical Welfare, Japan*
[b] *Adjunct instructor,Department of Health Informatics, Kawasaki University of Medical Welfare, Japan*
[c] *Asahigawasou Medical Welfare Academy School , Japan*

Abstract

A computer-based learning system called Electronic Patient Record (EPR) Laboratory has been developed for students to acquire knowledge and practical skills of EPR systems. The Laboratory is basically for self-learning. Among the subjects dealt with in the system is health information ethics. We consider this to be of the utmost importance for personnel involved in patient information handling. The variety of material on the subject has led to a problem in dealing with it in a methodical manner. In this paper, we present a conceptual model of health information ethics developed using UML to represent the semantics and the knowledge of the domain. Based on the model, we could represent the scope of health information ethics, give structure to the learning materials, and build a control mechanism for a test, fail and review cycle. We consider that the approach is applicable to other domains.

Keywords:

learning support system, health information ethics
conceptual model, UML, electronic patient record system

Introduction

In recent years, the introduction of electronic patient record (EPR) systems into hospitals has influenced the roles and responsibilities of personnel in charge of patient records significantly. At the Department of Health Informatics, Kawasaki University of Medical Welfare, we have about 80 student enrollments each year, and about a half of them aim to work for hospitals as patient record administrators or as health care information technologists in charge of health information systems. To assist students to prepare for EPR systems, we have been developing a computer-based learning system called Electronic Patient Record (EPR) Laboratory for over five years. The details about the EPR Laboratory are discussed in our previous paper (in Japanese) [1]. The system is designed to supplement conventional lectures on health information systems given as part of our undergraduate curriculum based on Recommendations on Health Informatics Education, 2001 by IMIA WG1. Using the Laboratory, the students learn the fundamentals of EPR systems including operational aspects and the subjects connected with patient information handling. It is basically meant for self-learning. When a student failed a test, the system shows which questions were not answered correctly. Then student must review the materials and try the test again. We call this a test, fail and review cycle.

Among the subjects dealt with in the EPR Laboratory is health information ethics. We consider it to be of the utmost importance for personnel involved in patient information handling. Although recently some textbooks on the subject have become available [2-4], health information ethics is still very new, relevant guidelines are updated or published yearly, and the domain is changing and evolving rapidly. There is a variety of material on a range of topics. The materials and tests were arranged sequentially in the Laboratory; however, topics are not necessarily sequentially ordered, and a sequential order is not appropriate for such a complex subject as health information ethics. It was harder to capture and represent the domain than in conventional subjects. Further, when a student failed a test, it was not necessarily possible to indicate automatically which part or all of materials that should be reviewed.

To overcome the problems, we developed a conceptual model of health information ethics using UML. The objectives were to represent the domain of health information ethics, to give structure to learning materials including documents (descriptive texts, published guidelines), exercises and tests, and to build a control mechanism for a test, fail and review cycle into the Laboratory. The definition of health information ethics itself is beyond the scope of the present study, since there is no commonly agreed definition. Instead, we present the developed conceptual model, describe how it is applied in the EPR Laboratory, and discuss the usefulness of the approach.

Materials and methods

Electronic patient record (EPR) laboratory

The present study makes use of the EPR Laboratory as a working bench. The EPR Laboratory was developed using Cache 5.0.11 (InterSystems). It is used in our Health Data Management Practice course for the third year students and delivered in a computer room with one we b server and 140 client PCs. The system is used only for educational purposes

and is not a full-fledged EPR system. The system is designed mainly 1for the students to learn the fundamentals of EPR systems through practices, and 2for instructors to prepare learning materials. EPR Laboratory aims to support the students to learn not only operational aspects but also healthcare services delivery and underlining information technology (IT). IT topics such as information security are covered in the EPR Laboratory in order that the students understand why they should acquire IT knowledge to deal with patient information. Details of IT topics such as SQL and XML are covered by other courses.

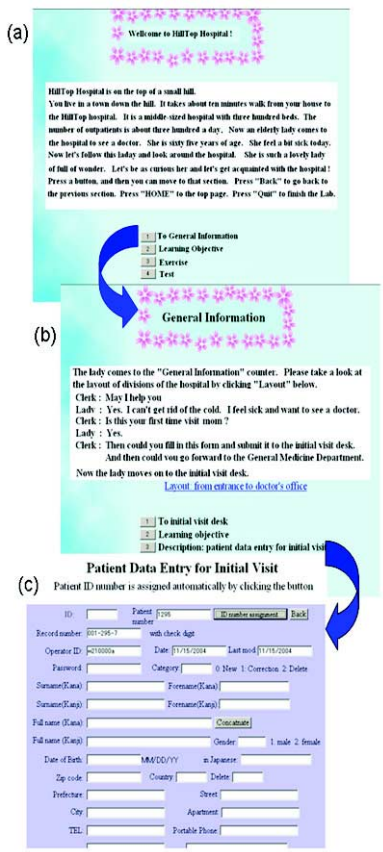


Figure 1 - EPR Labatory: a) The introductory section; b) Description about general information; c) Data entry for patient basic data

Learning materials are arranged in sections, and in each section, an objective(s), descriptive texts and exercises are given. Figure 1 shows the screens of the introductory section on the organization and functions of a hospital and patient flow. English messages in Figure 1 are prepared only for this paper, and in the actual system, all messages are shown in Japanese. The section starts with a fictitious story that an elderly lady comes to visit a hospital (Figure 1(a)). The descriptions continue as the lady moves from general information, initial visit desk, doctor's office, and so on. Figure 1(b) is the screen of general information. Figure 1(c) is the screen of the basic patient data entry, where the students practice operations. There are only two levels of access mode: the record administrator mode and the doctor mode. In the former mode, only limited operations are available including basic data entry, creation and registration of templates for data entry, semi-automatic coding of diagnoses (ICD-10), computation of hospital statistics, and so on. At certain point, there is a test, and the students may not move on to the next screen until they pass the test.

To assist instructors, facilities for preparing teaching materials and tests are provided. Questions on a test may be designed and edited using a textbox, check box, list box and so on, and the score for each question and the pass criteria of the test may be specified.

Conceptual modeling with UML

Modeling is a proven and well-accepted engineering technique in software engineering and it has been applied in a variety of health informatics studies [5-9]. In our study, a conceptual model was developed to solve the problems encountered during the development of the learning contents of health information ethics.

Results

We consider health information ethics a subject that everyone who has access to information systems with patient data should learn. Two levels of learning objectives may be distinguished: the first level for general users of health information systems, and the second level for those who are responsible for patient information handling. In the following, we will focus on the learning objectives of the second level; that is, the objectives for students who aim to be health information professionals.

A conceptual model of health information ethics

General outline of the model

Health information ethics covers both health care services (patient care) and medical and health research. The domain includes a wide range of topics and there are various types of relations among the topics. To represent the knowledge of the domain of health information ethics, we adopted the information modeling technique and developed a conceptual model using UML [10].Figure 2 shows the representation of a simplified version of the model. The model represents the scope of health information ethics covered in our current course including health ethics, information ethics, information security, EBM, and ethics for health information professionals. The conceptual model was developed to give structure to learning materials and to build a control mechanism into the test, fail and review cycle in the EPR Laboratory.

Terms and concepts

The model consists of classes and relationships among the classes. A class, rendered as a rectangle in UML, is used to model abstractions that are drawn from the domain health information ethics. In Figure 2, HealthInf and HealthEthics are examples of classes. We use a short name for a class because of space limitations. HealthInf, for example, is

short for health information. Each of the classes (abstractions) is a part of the vocabulary of the domain.

There are three kinds of relationships: dependencies, generalizations and associations. A 'dependency' is used to show one thing using another. It is rendered as a dashed line directed to the thing being depended on. In Figure 2, for example, there is a dependency from the class 'EBM' to the package 'Epidemiology', where a package (rendered as a tabbed folder) is the basic grouping with which we may organize a model. We use a package when the contents are relevant but are broad in scope and are taught outside the course. 'EBM' uses the basics of the subject 'Epidemiology' taught elsewhere. In software, a 'dependency' means a change in specification of one thing may affect another thing that uses it.

A 'generalization' is a relationship between a general thing (called the parent) and a more specific kind of that thing (the child). It is rendered as a solid line with a large open arrowhead pointing to the parent. In Figure 2, the class ClinicalInf (short for clinical information) is a child of HealthInf. A child inherits the properties of its parents, and it may also have its own properties in addition to those found in its parents. A class may have zero, one or more parents. A class that has no children is called a leaf class.

An 'association', rendered as a solid line connecting classes, is used to show structural relationships; that is, objects (to be described later) of one thing are connected to objects of another. An 'aggregation' is a special kind of association and is used to model a whole/part relationship. It is shown with a diamond at the end. For example, SecurityPolicy is a part of InfSecurity (Information Security) in Figure 2.

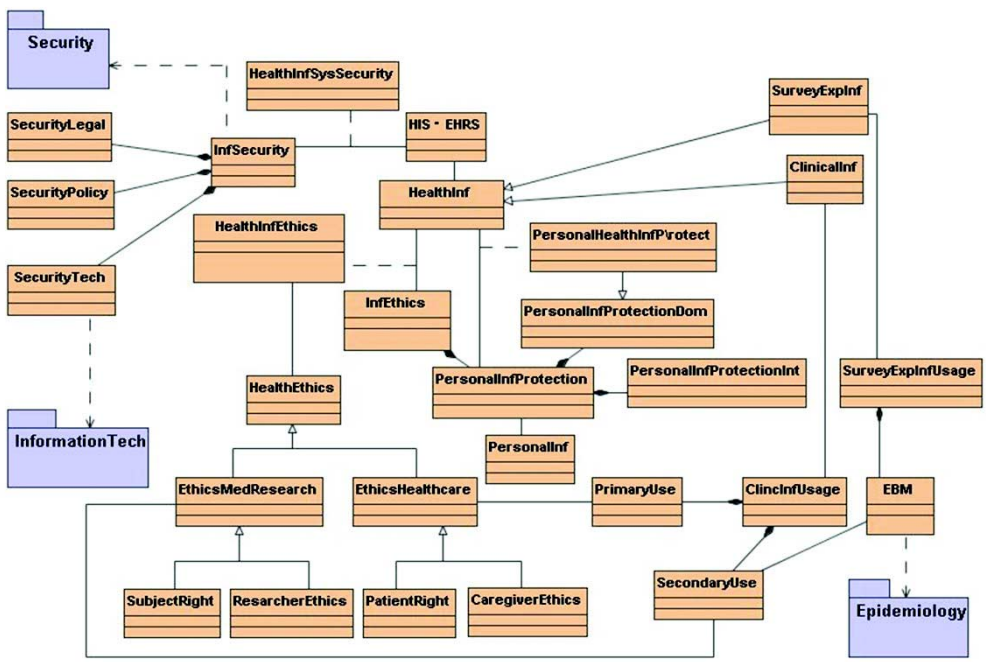


Figure 2 - A conceptual model of "Health Informatin Ethics" (simplified version)

Conceptual model in EPR laboratory

The conceptual model has been incorporated into the EPR Laboratory. Some of the implementation details are described below.

Objects in EPR laboratory

In our application, teaching materials (including documents, descriptive texts, exercises and tests) are physical things and are considered objects. Each object belongs to a class, and a class is a description of a set of objects that share the same attributes, relationships and semantics. There are also abstract classes (classes with no objects). Classes with objects may explicitly be called concrete classes. In our application, concrete classes are mostly leaf classes. In Figure 2, leaf classes are omitted because of space limitations. Below in ResearcherEthics for example, there are classes ResearcherEthicsGuideline, ResearcherEthicsText and ResearcherEthicsTest. The objects that belong to the class ResearcherEthicsGuideline are the guidelines including World Medical Association Declaration of Helsinki Ethical Principles for Medical Research Involving Human Subjects, and the important ethical guidelines on medical or epidemiological research published in Japan. There are dependency relations among the classes, e.g., from ResearcherEthicsTest to ResearcherEthicsText. This means some of the texts are used or referenced by a test, and changes in the texts may affect the test but not the reverse.

Attributes of classes in EPR laboratory

For concrete classes with learning materials as objects, the following attributes are defined, where attribute names are shown in italics. The attributes may be considered as metadata of learning materials. The materials may be documents (such as guidelines), tests and so on.

Class_name

a) title : the title of the material

b) description : the description of the material

c) responsible_entity : responsible entity for the material

d) year_month : year (and month) of publication or writing

e) depend_on : the title(s) of the material(s) that this material depends on

For a test, the following attributes are defined in addition to those defined in a) through e) above:

f) question_ids : identifications of the questions belonging to this test

g) pass_criteria : the pass criteria for this test

An example pass criteria for a test may be 'the sum of all questions is greater than 80 and Question 1 is greater than 5 and Question 7 is greater than 5'.

For a question comprising a given test, the following attributes are defined in addition to the attributes a) to e) above:

f) answer : answer(s) to this question

g) scoring_method : the method for scoring this question

Control of test-fail-review cycle

There are three levels of questions: fundamental, advanced, and applied. A fundamental question is such that an answer can be found easily in a learning material. A question is advanced if the answer may be obtained from multiple materials. A question is applied if an answer cannot be found directly from materials.

If a student fails in answering fundamental or advanced questions, they are guided to read the relevant materials that are indicated by the '*depend_on'* attribute. As an example, Figure 3 shows a fundamental question on personal data protection. The question is stored as an object of the class representing questions for the topic with the following attribute values:

a) title : question 1

b) description : Check if a student has an understanding of the outline of the OECD eight principles

c) responsible_entity : name of the composer of the question

d) year_month : 2005 Dec

e) depend_on : ResearcherEthicsGuideline. 'OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data'

f) answers : 1. e 2. b 3. a 4. g

g) scoring_method: type 3

The value of the attribute '*depend_on'* shows the material(s) that this question is dependent on; that is, the title of the guideline preceded by the class name. The guideline itself is stored as an object of the class 'ResearcherEthicsGuideline', with the following values of the attributes (shown only partly):

a) title : OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data

b) description : ……

For fundamental and advanced questions, answers are selected from a list of choices and the scores are computed automatically according to the rules specified by the attribute '*scoring_method'*.

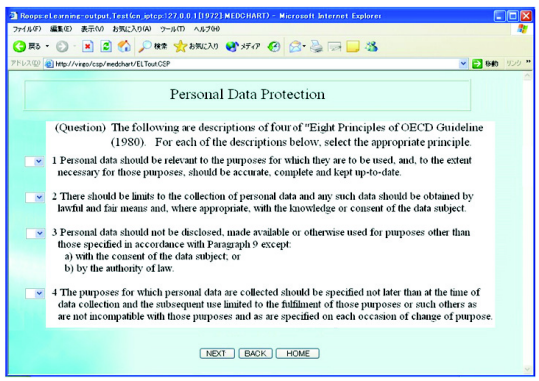


Figure 3 - A sample fundamental question on 'personal data protection'

Applied questions are given toward the end of the course. The questions are often taken from the real life problems that could be very complex. For this type of question, a student is asked to answer in writing in a textbox. The answer is sent to the instructor, and the instructor gives a score. A sample applied question is shown below.

A sample applied question:

> Assume that there was a large-scale accident, and a number of injured people, unconscious and unidentified, were carried into your hospital. Then you receive a phone call from a person saying that his relative might have been involved in the accident. He would like to know whether the named person is among the injured people brought into your hospital. Is it appropriate for you to answer his question over the phone? Discuss why you consider appropriate or not appropriate.
>
> 1. appropriate 2. not appropriate

The applied questions on personal data protection are taken from Q & A for The Guideline on Healthcare and Welfare Entities Handling of Personal Information issued by the Japanese Ministry of Health, Labour and Welfare.

Discussion and conclusion

The EPR Laboratory has been developed for over five years to assist students to prepare for Electronic Patient Record (EPR) systems. The system is meant to supplement conventional lectures. Among the subjects covered in the EPR Laboratory, is health information ethics. For a newly emerging concept such as health information ethics, it was necessary to have a method to represent the knowledge of the domain and organize learning materials in a systematic way. Further, it was also necessary to build a mechanism to tell the students about their weak points when they failed a test so they could review materials more appropriately.

A conceptual model of health information ethics was developed to organize the contents of health information ethics in a way that expresses the semantics of the domain more visibly. Modeling is a proven and well-accepted engineering technique in software engineering, and many aspects of modeling techniques for software apply to our application as well.

The contents of health information ethics in the EPR Laboratory are intended to educate students who aim to be health information professionals. We consider that ability for making decision on real problems of health information ethics may only be acquired through practice in health care settings but, at the same time, we consider students should acquire the fundamental principles before they are involved in the real-life decision-making. The Laboratory may also be used to educate students aiming for to be health care professionals, or for training of health care professionals.

Incorporation of the model into the EPR Laboratory provides a template for storing learning materials in an orderly way, and builds a control mechanism for a test, fail and review cycle. It also provides a means to thoroughly examine the necessity and suitability of the contents. We consider that the approach is applicable to other domains.

References

[1] Watanabe K., Yamamoto K., Okada M., and Takaue R. A computer aided instruction for electronic patient records in a hospital. Japan Journal of Medical Informatics 2006: 25(4): 249-56. (published in March 2006 in Japanese)

[2] Chapman AR. (Ed). Health Care and Information Ethics: Protecting Fundamental Human Rights. Kansas City: Sheed & Ward, 1997.

[3] Anderson J.G. and Goodman KW. Ethics and Information Technology: A Case-Based Approach to a Health Care System in Transition. New York: Springer, 2002.

[4] Harman LB. Ethical Challenges in the Management of Health Information. Maryland: Aspen, 2001.

[5] Okada M., Hashimoto H., and Ohida T. Domain analysis and modeling to improve comparability of health statistics. In: Proc. MEDINFO2001, Patel V. et al., eds. Amsterdam: IOS Press, 2001; pp.1374-78.

[6] Burek P., Hoehndorf R., Loebe F., Visagie J., Herre H., and Kelso J. A top-level ontology of functions and its application in the Open Biomedical Ontologies. Bioinformatics 2006: 22(14): e66-73.

[7] Garde S., Knaup P., and Qumquad H.R. A UML-based approach for remodeling of legacy systems in health care. Int J Med Inform 2003: 70(2-3): 183-94.

[8] Ruggieri A.P., Elkin P.L., Solbrig H., and Chute C.G. Expression of a domain ontology model in unified modeling language for the World Health Organization International classification of impairment, disability, and handicap, version 2. Proc AMIA Symp 2001: 573-7.

[9] Bornberg-Bauer E. and Paton N.W. Conceptual data modelling for bioinformatics. Brief Bioinform 2002: 3(2):166-80.

[10] Booch G., Rumbaugh J., and Jacobson I. The Unified Modeling Language User Guide, 2nd Edition. Reading: Addison-Wesley, 2005.

Address for correspondence

288 Matsushima,
Kurashiki City,
Okayama 701-0193,
Japan

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

PDA-based Informatics Strategies for Tobacco Use Screening and Smoking Cessation Management: A Case Study

Suzanne Bakken[a,b], W. Dan Roberts[c], Elizabeth Chen[b], Joann Dilone[a], Nam-Ju Lee[a], Eneida Mendonca[b], Marianthi Markatou[d]

[a] *Columbia University School of Nursing, New York, New York*
[b] *Department of Biomedical Informatics, Columbia University, New York, New York*
[c] *Adelphi University, Garden City, New York*[d] *Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, New York*

Abstract

The purpose of this case study is to describe three incremental personal digital PDA-based informatics strategies aimed at improving screening for tobacco use and guideline-based tobacco cessation management: 1) PDA clinical log with tobacco cessation diagnoses and plan of care options, 2) PDA decision support system, and 3) PDA decision support system with infobuttons - context-specific links to the National Cancer Institute's Cancer Information Services tobacco cessation information. These strategies were implemented within the context of an evidence-based advanced practice nurse curriculum at the Columbia University School of Nursing.

Keywords:

tobacco cessation, decision support, nursing informatics

Introduction

Tobacco use is a significant global issue. Cigarette smoking is the leading cause of lung cancer and mortality related to lung cancer and has been implicated in other cancers such as larynx, oral cavity and pharynx, esophagus, bladder, stomach, pancreatic, liver, renal cell and renal pelvis, cervix, and acute myeloid leukemia [1, 2]. Tobacco use contributes to 430,000 deaths per year in the U.S., which according to Healthy People 2010 represents 5 million years of potential life lost [3]. Healthy People 2010 further reports that fifty billion dollars per year of direct medical costs are a result of cigarette smoking. In the U.S. as in other countries, the goal of reducing tobacco use remains an urgent necessity for improving both quality and length of life. Meeting Healthy People 2010 goals related to tobacco cessation is estimated to prevent an addition 7.1 million premature deaths [2].

The purpose of this case study is to describe three incremental personal digital PDA-based informatics strategies aimed at improving screening for tobacco use and guideline-based tobacco cessation management: 1) PDA clinical log with tobacco cessation diagnoses and plan of care options, 2) PDA decision support system, and 3) PDA decision support system with infobuttons - context-specific links to the National Cancer Institute's Cancer Information Services tobacco cessation information. We also provide descriptive data regarding use of the first and second strategies.

These strategies were implemented within the context of an evidence-based advanced practice nurse (APN) curriculum at the Columbia University School of Nursing. The first strategy was developed to serve as a control for the second strategy in a randomized controlled trial that is currently in progress. The third strategy is in the design phase for a second randomized controlled trial.

Tobacco cessation and nursing practice

Although tobacco use prevention and smoking cessation counseling are within the purview of nursing practice, actual or perceived barriers frequently deter nurses from undertaking such interventions. McCarty et al. reported reasons given by nurses for not advising patients about smoking cessation: lack of a nursing time, lack of knowledge related counseling techniques, and reluctance to ask a patient to give up a coping mechanism during a hospitalization were given as reasons for nurses not to advise patients about smoking cessation [4]. They also found that nurses often felt counseling was ineffective. Current smokers in this study group were less likely to give smoking cessation advice. In another study, McCarty et al. found that only 50% of patients identified as smokers received smoking cessation advice [5]. Similarly, Gomm et al. reported nurses' attitudinal barriers of 'not my place', 'they'll quit if they want to', and not wanting to add to the patient's 'stress' [6]. In a study of pregnant women, midwifery assessment of smoking status and providing advice to pregnant women was accepted as part of the routine care [7]. Further, the women in the study didn't believe that assessing smoking status or providing advice altered the relationship between them and their midwife. Therefore, the nurse's perception may not be aligned with the perception of the patient about smoking cessation advice.

According to the U.S. Surgeon General's report [8], approximately 36 million American smokers will have at

least one outpatient visit each year but only about half of these smokers will receive smoking cessation advice at their visit. This is consistent with data from the Robert Wood Johnson Foundation that suggests that almost all obstetricians report asking about tobacco use in pregnancy, but fewer than half go on to discuss cessation strategies and offer self-help materials [9]. Smoking cessation advice and treatment is within the purview of the APN. Consequently it is important to design strategies that support guideline-based care and to integrate these into the curriculum so that APN students can address this important health concern during and after their clinical training.

Decision support and guideline-based care

For more than two decades computers have assisted in the provision of reminders to clinicians regarding standardized protocols or guidelines [10-12]. A number of RCTs have demonstrated that computer-based reminders decrease errors of omission [10, 13, 14] and increase compliance with preventive care guidelines [15, 16]. Several recent systematic reviews also suggest that such systems impact clinician adherence to guidelines [17-19].

Additional studies support the potential for decision support to affect clinician behavior related to smoking cessation. A study of physicians in Vermont revealed that even though they were familiar with the U.S. Public Health Service guideline [8] on tobacco use and tobacco dependent treatment and had positive attitudes toward it, they were unfamiliar with smoking cessation resources [20]. The authors concluded that interventions to improve clinician adherence to the guideline should address the inaccurate perception that smokers are unreceptive to counseling and increase knowledge of smoking cessation resources and that a decision support system could meet these needs. An randomized controlled trial of a multifaceted intervention that included computer-generated prompts for care provision by clinic staff found that patients in the experimental group were more likely to report receiving smoking cessation advice by nursing and anesthesia staff and to receive nicotine replacement if nicotine dependent [21].

PDA clinical log at the Columbia Univ. School of Nursing

The PDA clinical log has been in use for more than four years and was designed to serve multiple purposes. These include: student documentation of clinical encounters using standardized nursing terminology and other health-care-related coding systems; use of benchmarking reports to learn to critically examine one's practice over time; and faculty monitoring of student performance. Key steps that were required to create a system that supported these multiple purposes were: design of the system architecture, selection of data elements and standardized terminologies for the data elements, design and implementation of the user interface, design and implementation of the database and knowledge base, and design and implementation of reports. These details have been described in detail elsewhere [22, 23].

APN students enter data using Palm OS devices and synchronize with a central database through Ethernet cradles, WiFi, Internet, or cellular network depending upon their location and type of device (e.g., Palm TX, Treo 700). Faculty and student reports are generated every two weeks.

Use of the PDA Clinical Log is a required part of the APN clinical curriculum. APN student participation in the randomized controlled trial that is testing the affect of a PDA support system on clinical care processes is optional.

Infobuttons and infobutton manager

Infobuttons are context-specific links between a clinical information system and web-based information resources that attempt to anticipate users' information needs and to automatically satisfy those needs [24-27]. Infobuttons are built upon three premises: high-quality, web-based resources exist and are continuously updated by their developers; clinical information systems should access existing web-based information resources, rather than develop and maintain content; and context-specific links within clinical information systems can provide more efficient access to web-based information resources. Infobuttons are currently deployed within two clinical information systems at Columbia University Medical Center: WebCIS and Eclipsys, as well as at two other institutions. In addition, they are accessible via the PDA in an experimental system called PalmCIS [28]. Cimino has recently described the use and positive influence of infobuttons on clinical care [29].

Materials and methods

PDA clinical log with tobacco cessation diagnoses and plan of care options

To create an appropriate application to serve as the control for the PDA decision support system, we needed to make sure it was possible to document all aspects of guideline-based care, i.e., the diagnosis of tobacco dependence and the associated plan of care items for Diagnostics, Procedures, Prescriptions, Patient Teaching and Counselling, and Referrals. First, the U.S. Public Health Service guideline was decomposed into its constituent elements by an expert in tobacco cessation and then reviewed by an additional expert in tobacco cessation and an expert in guideline representation [30]. Then, we represented the diagnostic and intervention concepts with codes from existing terminologies or created new codes in instances where existing terminologies were insufficient to represent the concepts. Next, the guideline-related terms were then organized into existing lists of terms in the menus for diagnoses and the five categories in the plan of care.

PDA decision support system

The PDA decision support system for tobacco screening and smoking cessation management provides decision support for a short clinical intervention based upon the Ask, Advise, Asses, Assess, and Arrange smoking cessation model [8, 30, 31]. As shown in Figure 1, there are two

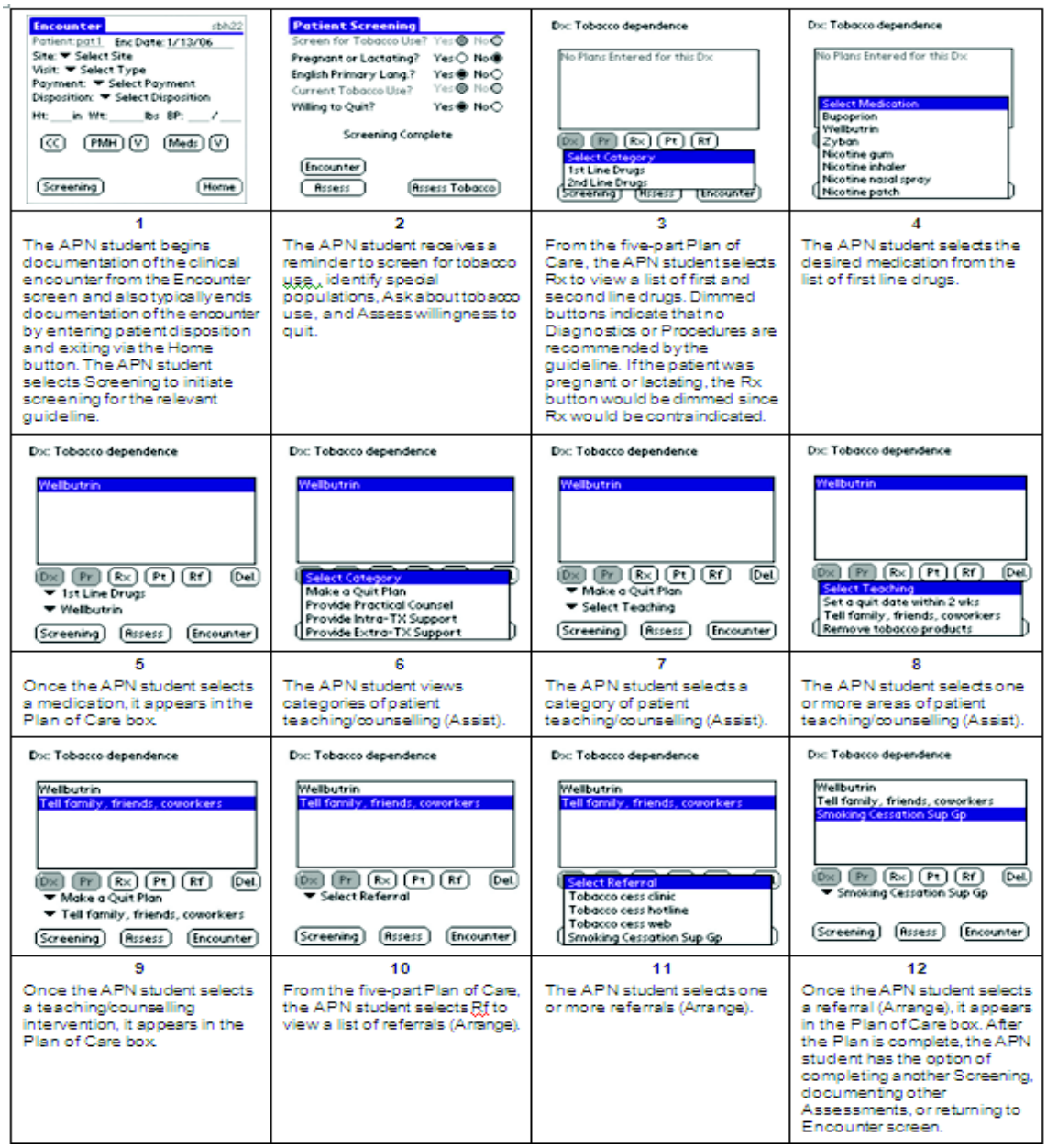


Figure 1 - PDA decision support system for tobacco cessation

aspects to the decision support: the student is prompted to screen and is presented with a guideline-based documentaion template for a plan of care that is tailored according to the smoker's willingness to quit and whether or not they are pregnant or lactating.

PDA decision support system with Infobuttons

Infobuttons are enabled though an Infobutton Manager, which provides a standardized interface for matching user contexts to information resources. This is done through six different methods for accessing the information resources: simple link, concept-based link, simple search, concept-based search, intelligent agent, and calculator [24].

For the PDA decision support system with infobuttons, our design is based upon the use of simple links and concept-based links to access the National Cancer Institute's Cancer Information Service tobacco cessation information within the Patient Teaching and Counselling and Referrals sections of the guideline-based document template. For example, if an APN student wishes to access patient edu-

cation materials in Spanish about making a Quit Plan, she will select the infobutton associated with that intervention and through cellular telephone technology access the Cancer Information Service resource of relevance.

Because the Cancer Information Service resources are not designed to optimize viewing on small devices such as PDAs and cellular telephones, we must also develop and implement an approach for displaying this information in a manner suitable for these small devices.

Results

Descriptive data are reported for Fall semester 2005 (PDA clinical log with tobacco cessation diagnoses and plan of care options) and for Fall 2006 (PDA decision support system) so that they reflect similar points in the APN curriculum. The third strategy is still in the design phase.

PDA clinical log with tobacco cessation diagnoses and plan of care options

In Fall of 2005 when all APN students received the PDA clinical log that included tobacco cessation diagnoses and plan of care options, 1132 encounters were documented using the PDA clinical log for persons more than 9 years old. Only one encounter included a tobacco dependence diagnosis. The associated plan of care included teaching and counselling related to “Make a Quit Plan”.

PDA decision support system

In Fall of 2006, 13 students randomized to received decision support for tobacco cessation entered 150 encounters in which they received a reminder to screen. Patients were screened in 64% (n=96) of the encounters with reminders. Fifteen of the 96 screened were identified as current smokers and 8 indicated that they were willing to quit. However, only one tobacco-related plan of care was documented in the guideline-based documentation template.

The majority of those screened were female (75%) and either Hispanic (42%) or Black (27%). Forty-four percent were between 19-35 years of age. Of those screened, 30% were pregnant or lactating and 42% did not speak English as their primary language. Among those identified as currently smoking, 33% were pregnant or lactating and 73% did not speak English as their primary language.

Discussion

The screening rate for those APN students who received the PDA decision support reminder to screen could be viewed as encouraging, but this only resulted in an associated plan of care in one of eight patients identified as willing to quit smoking. Further study is needed to determine if this is an artifact of the documentation process or an actual quality of care issue. Infobuttons have the potential to provide additional decision support through provision of access to Cancer Information Service resources. The effect of the PDA decision support system on screening rates and adherence to guideline recommendations is currently being tested in a randomized controlled trial in which the control groups receive PDA decision support for either depression or obesity management. A second randomized controlled trial will compare the PDA decision support system and the PDA decision support system with infobuttons on use of Cancer Information Service resources by APN students and the patients they refer to the resources.

Conclusion

There has been little work focused on decision support for APN practice or specifically for tobacco screening and smoking cessation management. Our case study describes our incremental approach to designing PDA-based informatics strategies for screening and management of tobacco dependence within the context of an evidence-based APN curriculum. It is vital that informatics methods and information technologies be conscientiously applied to address this significant health issue from both the clinician and consumer perspectives. Our hope is that by integrating these approaches into the curriculum, the benefits will affect not only the patients the nurses care for during their APN education but also the patients they will care for after graduation through application of such tools into their APN practice.

Acknowledgements

This research is funded by the National Institute of Nursing Research (R01 NR008903) and the National Cancer Institute (R21 126325). The authors thank the other members of the research teams for their contributions.

References

[1] Sasco AJ, Secretan MB, Straif K. Tobacco smoking and cancer: a brief review of recent epidemiological evidence. Lung Cancer 2004;45 Suppl 2:S3-9.

[2] The Health Consequences of Smoking: A Report of the Surgeon General. Washington, DC: Department of Health and Human Services May 27, 2004.

[3] Office of Disease Prevention and Health Promotion. Healthy People 2010: Understanding and Improving Health. Washington, DC: Office of Disease Prevention and Health Promotion (ODPHP), United States Department of Health and Human Services; 2000.

[4] MCarty M, Zander KM, Hennrikus D, Lando H. Barriers among nurses to providing smoking cessation advice to hospitalized smokers. Am J Health Promot 2001;16(2):85-87.

[5] McCarty M, Hennrikus D, Lando H, Vessey J. Nurses' attitudes concerning the delivery of brief cessation advice to hospitalized smokers. Prev Med 2001;33(6):674-681.

[6] Gomm M, Lincoln P, Egeland P, Rosenberg M. Helping hospitalised clients quit smoking: a study of rural nursing practice and barriers. Aust J Rural Health 2002;10(1):26-32.

[7] McCurry N, Thompson K, Parahoo K, O'Doherty E, O'Doherty A. Pregnant women's perceptions of the implementation of smoking cessation advice. Health Educ J 2002;61(1):20-31.

[8] Centers for Disease Control and Prevention. Surgeon General's Report: Reducing Tobacco Use. Atlanta, GA: Centers for Disease Control and Prevention; 2000.

[9] The Robert Wood Johnson Foundation. Substance Abuse: The Nation's Number One Health Problem. Princeton, NJ: The Robert Wood Johnson Foundation; 2001.

[10] McDonald CJ. Protocol-based reminders, the quality of care, and the non-perfectability of man. New Engl J Med 1976;295:1351-1355.
[11] McDonald CJ, Hui SL, Smith DM, Tierney WM, Cohen SJ, Weingerger M, et al. Reminders to physicians from an introspective computer medical record: A two-year randomized trial. Ann Intern Med 1984;100:130-138.
[12] Barnett GO, Winickoff RN, Morgan MM, Zielstorff RD. A computer-based monitoring system for follow-up of elevated blood pressure. Med Care 1983;21(4):400-409.
[13] Burack RC, Gimotty PA, George J, Stengle W, Warbasse L, Moncrease A. Promoting screening mammography in inner-city settings: A randomized controlled trial of computerized reminders as a component of a program to facilitate mammography. Med Care 1994;32(6):609-624.
[14] Overhage JM, Tierney WM, Zhou X, McDonald CJ. A randomized trial of "corollary orders" to prevent errors of omission. J Am Med Inform Assoc 1997;4(5):364-375.
[15] McPhee SJ, Bird JA, Fordham D, Rodnick JE, Osborn EH. Promoting cancer prevention activities by primary care physicians: Results of a randomized, controlled trial. JAMA 1991;266(4):538-544.
[16] McDonald C, Hui S, Tierney WM. Effects of computer reminders for influenza vaccination on morbidity during influenza epidemics. MD Comput 1992;9:304-312.
[17] Johnston ME, Langston KB, Haynes RB, Mathieu A. Effects of computer-based clinical decision support systems on clinician performance and patient outcome: A critical appraisal of research. Ann Intern Med 1994;120(2):135-142.
[18] Shea S, DuMouchel W, Bahamonde L. A meta-analysis of 16 randomized controlled trials to evaluate computer-based clinical reminder systems for preventive care in the ambulatory setting. J Am Med Inform Assoc 1996;3(6):399-409.
[19] Zielstorff RD. On line practice guidelines: State of our knowledge and a look to the future. J Am Med Inform Assoc 1998;5(3):227-36.
[20] Marcy TW, Skelly J, Shiffman RN, Flynn BS. Facilitating adherence to the tobacco use treatment guideline with computer-mediated decision support systems: physician and clinic office manager perspectives. Prev Med 2005;41(2):479-87.
[21] Wolfenden L, Wiggers J, Knight J, Campbell E, Spigelman A, Kerridge R, et al. Increasing smoking cessation care in a preoperative clinic: a randomized controlled trial. Prev Med 2005;41(1):284-90.
[22] Bakken S, Cook S, Curtis L, Desjardins K, Hyun S, Jenkins M, et al. Promoting patient safety through informatics-based nursing education. Int J Med Inf 2004;73(7-8):581-589.
[23] Klein WT, Bakken S. Design and Implementation of a Student Clinical Log Database and Knowledge Base. In: NI 2003; Rio de Janeiro, Brazil; 2003.
[24] Cimino J, Li J, Allen M, Currie L, Graham, M, Janetzki V, et al. Practical considerations for exploiting the world wide web to create infobuttons. Medinfo. 2004 2004(277-281).
[25] Allen M, Currie LM, Graham M, Bakken S, Patel VL, Cimino JJ. The classification of clinicians' information needs while using a clinical information system. AMIA Annu Symp Proc 2003:26-30.
[26] Currie LM, Graham M, Allen M, Bakken S, Patel V, Cimino JJ. Clinical information needs in context: an observational study of clinicians while using a clinical information system. AMIA Annu Symp Proc 2003:190-4.
[27] Lei J, Chen ES, Stetson PD, McKnight LK, Mendonca EA, Cimino JJ. Development of infobuttons in a wireless environment. AMIA Annu Symp Proc 2003:906.
[28] Chen ES, Mendonca EA, McKnight LK, Stetson PD, Lei J, Cimino JJ. PalmCIS: A wireless handheld application for satisfying clinician information needs. J Am Med Inform Assoc 2004;11(1):19-28.
[29] Cimino JJ. Use, usability, usefulness, and impact of an Infobutton Manager. Proc American Medical Informatics Association Symp 2006:151-155.
[30] Roberts WD, Patel VL, Stone PW, Bakken S. Knowledge content of advanced practice nurse and physician experts: A cognitive evaluation of clinical practice guideline comprehension. In: 9th International Congress on Nursing Informatics; 2006; Seoul, Korea; 2006.
[31] Bakken S, Jenkins M, Choi J, Hyun S, John R, Joyce M, et al. Usefulness of a Personal Digital Assistant-based advanced practice nursing student clinical log: Faculty stakeholder exemplars. In: 9th International Congress on Nursing Informatics; 2006; Seoul, Korea; 2006.

Address for correspondence

Suzanne Bakken, RN, DNSc
Columbia University
Mailbox 6
New York, NY 10032
e-mail: suzanne.bakken@dbmi.columbia.edu

This page intentionally left blank

Chapter 12.

Poster Contributions

Selected for
Best Poster Awards 2007

This page intentionally left blank

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Needs Assessment for the Computer-interpretable Hypertension Guideline at Public Health Centers in Korea

EunJung Lee[a], SoYoung Kim[b], InSook Cho[c], JiHyun Kim[a], JaeHo Lee[d], Yoon Kim[b]

[a] *Center for Interoperable EHR, Korea*
[b] *Department of Health Policy and Management, Seoul National University College of Medicine, Korea*
[c] *Department of Nursing, Inha University College of Medicine, Korea*
[d] *Department of Emergency Medicine, University of Ulsan College of Medicine, Korea*

Abstract

Computer-interpretable hypertension guidelines can make for the improvement of blood pressure control rate by assisting clinicians at the point-of-service to behave according to the evidence-based guidelines. We surveyed 117 public health centers in Korea to evaluate current hypertension management status with clinicians' performance, and to analyze the needs for the computer-interpretable hypertension guidelines. Hypertension control rate was 57%, while clinicians overestimated it as 79.5%. 60.4% of patients were treated with 2 or more antihypertensive medications, and most frequently used drug class was calcium channel blocker. Inappropriate prescription rate of contraindicated patients was 2.6%. Two-thirds of clinicians agreed to implement the computer-interpretable hypertension guideline. Implementation of computer-interpretable hypertension guideline is considered as a way to improve the hypertension control rate and to reduce the inappropriateness of the therapeutic choice.

Keywords:

needs assessment, hypertension, guideline, practice guideline

Introduction

For the prevention of cardiovascular diseases, hypertension management takes important portion. Yet world-wide hypertension control rate is known to be less than one-third, especially in Korea, it is about 17%. Computer-interpretable guideline is suggested as an intervention tool for the quality improvement of patient care as modifying clinicians' behavior. This is a preliminary study for developing computer-interpretable hypertension guideline, which is designed to improve hypertension control rate in primary health care. We intended to analyze current hypertension management status, clinicians' performance and needs for the CIGs.

Materials and methods

We surveyed 117 public health centers and 154 clinicians in Gyeonggi province. We extracted medical records of hypertension patients from the centers' clinical database during a 10-month period. Data included blood pressure, medication, comorbid disease, and others. Questionnaire was answered by the clinicians to assess their adherence to hypertension guidelines and needs for the CIG program.

Results

38,474 patients' data and 41 clinicians' survey was analyzed. 61% of clinicians agreed to implement the computer-interpretable hypertension guideline. Hypertension control rate was 57%. Clinicians overestimated as 79.5% of patients had controlled blood pressure. Over 60% of patients were treated with 2 or more antihypertensive medications, and calcium channel blocker was the most frequently prescribed. Based on JNC7 report, inappropriate prescriptions, such as monotherapy for the stage hypertensive patient, beta blocker only for the diabetes patient, combination of beta blocker and non-dihydropyridine calcium channel blocker, thiazide diuretics for the gout patient, and beta blocker for the asthma patient, were 2.6% (N=1001).

Conclusion

Clinicians overestimated the proportion of their patients with controlled blood pressure. Implementation of computer-interpretable hypertension guideline is considered as a way to reduce the inappropriateness of the therapeutic

Address for correspondence

EunJung Lee, M. D. toro0117@snu.ac.kr

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Development of a Personal Medical Recorder on a Cell Phone

Akihiro Takeuchi[a], Katsura Kobayashi[b], Noritaka Mamorita[b], Noriaki Ikeda[a]

[a] *Departments of Medical Informatics, School of Allied Health Sciences, Kitasato University, Japan*
[b] *Graduate School of Medical Sciences, Kitasato University, Japan*

Abstract

Paper medical records have effectively been used in chronic diseases without information technology. To facilitate self-control in hemodialysis and observe a patient's condition continuously, we developed a mobile phone-based personal medical recorder for patients suffering chronic renal failure. The application is based on Java2 Micro Edition and operates like a scheduler. The application stores laboratory data, such as BP, BUN, creatinine, HbA1c, etc., and other pertinent clinical comments into memory on a cell phone. The application can also customize, add or delete items (laboratory data, medications, questions, etc.). Detailed graphic displays of the data are shown. The data can also be sent to a PC with infrared communications. In a usage trial, patients were favorably receptive about this application and indicated that they wanted to continue using it.

Keywords:

cell phone, medical records, infrared transmitter

Introduction

Using a diary, a chronically ill person can gain an overview of his medical situation [1]. Although patients with chronic renal failure can maintain an adequate daily nutritional intake, they might also be able to recognize their own pathophysiological state and be aware of the settings for their own hemodialysis. Although data collection was more reliable with a palmtop computer than with paper diaries [2], there have been no reports presenting a PDA/cell phone-based clinical recorder for chronic hemodialysis other than for dietary monitoring [3]. Therefore, we developed a cell phone-based medical recorder for the self-management of chronic diseases.

Materials and methods

An application was designed as a self-controllable tool for chronic diseases. It consists of a calendar canvas, dataentry canvas, plot canvas, memory control module and an infrared module (Figure 1, left). The plot canvas shows graphs of clinical and laboratory data, such as BP, BUN, creatinine, HbA1c, etc. These data are stored in text form in the memory of a cell phone (maximum 200 kB). The data size of this schedule is estimated to be about 100 bytes. The monthly or daily data are sent to a PC by an infrared module.

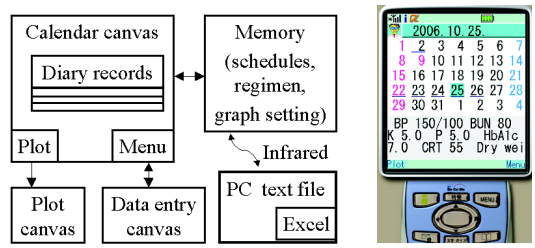


Figure 1 – Overview of the application and a calendar canvas

Results

The application was based on Java2 Micro Edition and i-Appli Tool (NTT DoCoMo, Inc.). The application (medData.jar 100 kB) operates like a scheduler. (Figure 1, right). All of the data, BP, laboratory data, medications, etc. on each day are listed on the data entry canvas. The patient interactively and personally types in numeric data and/or comments. Detailed graphic displays of data are shown sequentially on a plot canvas in years, months, weeks, or days. Color and scale customization of the plot canvas is possible for each set of laboratory data. Customizing items (laboratory data, medications, questions, etc.) is done on the data entry canvas. The application was adapted for several patients to use at the same time in one trial. The patients were favorably receptive about this application and said that they would like to continue using it. This personal medical recorder based on a cell phone could be useful as a self-management tool for patients with chronic diseases.

References

[1] Baert H, Gielen P, and Smet M. A health diary for the chronically ill. WHO Reg Publ Eur Ser 1992: 44: 328-31.
[2] Hyland ME, Kenyon CA, Allen R, et al. Diary keeping in asthma: comparison of written and electronic methods. *BMJ* 1993: 306: 487-9.
[3] Sevick MA, Piraino B, Sereika S, et al. A preliminary study of PDA-based dietary self-monitoring in hemodialysis patients. J Ren Nutr 2005: 15: 304-11.

Address for correspondence

take@kitasato-u.ac.jp, Kitasato, Sagamihara, Japan

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Implementing and Evaluating a Laboratory Information System to Optimize the Treatment of Tuberculosis Patients in Peru

Joaquin A. Blaya[a,b], Sonya S. Shin[b,c], Martin JA Yagui[d], Luis Asencios[d], Javier Vargas[d], Carmen Suares[e], Gloria Yale[f], Hamish SF Fraser[b,c]

[a] *Harvard-MIT Division of Health Sciences & Technology, USA,* [b] *Partners In Health, USA,* [c] *Division of Social Medicine &Health Inequalities, Brigham & Women's Hospital, USA,* [d]*Instituto Nacional de Salud, Perú,* [e]*Dirección de Salud IV Lima Este, Perú,* [f]*Dirección de Salud V Lima Ciudad, Perú*

Abstract and objective

Multi-drug resistant tuberculosis patients in resource-poor settings experience large delays in starting appropriate treatment and may not be monitored appropriately due to an overburdened health care system, communication delays, and missing or error-prone data. A web-based laboratory information system "e-Chasqui" has been implemented in Peru to alleviate these problems by improving the timeliness and quality of laboratory data. It has been deployed in the national TB laboratory, two regional laboratories and twelve health centers. High user satisfaction and heavy use has led to e-Chasqui is being expanded to more institutions. A study is being performed to measure its impact and generalizability.

Keywords:

clinical laboratory information system, computerized; evaluation studies; developing countries

Introduction

Treatment for MDR-TB in Peru is often delayed by over three months after initial presentation.(1) These potentially dangerous delays occur because of the long test processing time, cumbersome collection and communication procedures, and loss of specimens and test results. Similar problems are prevalent in other settings including South Africa.(2). An electronic information system can improve the handling and communication of these data between institutions. Decreasing treatment delays and ensuring an appropriate drug regimen should improve outcomes and reduce transmission.

Methods

Partners In Health has developed a web-based medical record system (PIH-EMR)(3) to support the treatment of TB, with data on 15523 patients. We created a web-based laboratory information system, "e-Chasqui" to connect laboratories to health centers to reduce delays and facilitate communication and analysis. e-Chasqui includes tools to improve data quality, notify health centers of new results, alert physicians of high-risk patients and create laboratory reports. Here we report on the system's implementation, use, and preliminary results from its evaluation.

Results

e-Chasqui has been implemented in the national reference laboratory, two of five regional laboratories in Lima and twelve health centers. Since its implementation in November, 2005, 19900 smear, 21196 culture and 3076 drug sensitivity test results have been entered. In 2006 all health centers have viewed 100% of their results online. Due to user satisfaction and heavy use, public officials have asked to expand the system to 3 other laboratories and over 10 other health centers.

Conclusions

This experience demonstrates the possible benefits of implementing a web-based laboratory information system in a low resource setting. A prospective randomized evaluation is being performed to measure its impact on delays, errors, and quality of care, including time to prescribe an effective drug regimen.

References

1. Yagui M, Perales MT, Asencios L, et al. Timely diagnosis of MDR-TB under program conditions: is rapid drug susceptibility testing sufficient? Int J Tuberc Lung Dis. 2006 Aug;10(8):838-43.
2. Hurtado R. Personal Communication. 2006.
3. Fraser H, Jazayeri D, Mitnick C, Mukherjee J, Bayona J. Informatics Tools To Monitor Progress And Outcomes Of Patients With Drug Resistant Tuberculosis In Peru. Proc AMIA Symp. 2002:270-4.

Address for correspondence

HSF Fraser Hamish_Fraser@hms.harvard.edu

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Risk analysis – A Tool for IT Development and Patient Safety
A Comparative Study of Weaknesses Before and After Implementation of a Health Care System in the County Council of Ostergotland, Sweden

Annica Öhrn[a], Gunilla Eriksson[b]

[a] *Patient safety office of county council of Ostergotland, Sweden*
[b] *Laboratory medicine unit of county council of Ostergotland, Sweden*

Abstract

There is a lack of tools to secure patient safety considering information security and health care systems. Risk analysis can be used as a systematic method for identifying risks before a medical error occurs. In healthcare is this type of analysis uncommon use. An indicator for this need of systematic proactive work is the amount of adverse advents that are reported in the deviation system .In this comparative study of a health care system during 2004 and 2006, one of the results is an reduction in the numbers of risks due to technique & equipment and in the area of training & competence. This study gives an indication of that risk analysis helps the health care organisation to develop and manage routines for reducing possible risks.

Keyword:

information system, safety management, risk management/methods, medical errors, systems analysis, quality of health care

Introduction

In the county council in Ostergotland, Sweden has a supporting system for requests and laboratory reports, called LR, and been implemented during the last year and a half. A risk analysis has been performed before and after implementation and is planned to continue regulary for the coming years. The aim of this study was to perform a comparative risk analysis for a recently implemented health care system.

Methods

Risk analysis has been performed before and after the implementation of the system. Risk analysis is a systematic tool, including: process mapping, risk identification, determination of the severity and probability of each risks and action plans. The analysis has a patient safety perspective and has been performed with a trained risk analysis team including members from different roles/specialities. The risks were categorised in areas of six potential sources; technique & equipment, rules/policies/procedures, environment, training & competence, barriers and communications & information.

Results

The risk analysis shows that the total risk points are lower during 2006 for the areas technical solutions and procedures compared to the points of the same areas during 2004.

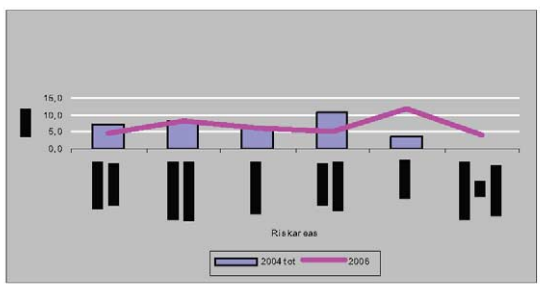


Figure 1 – Risk profile, comparison 2004-2006 LR

Discussions

The risk analysis performed demonstrated differences between the results obtained prior and after the implementation. This was especially true within the sources; technique & equipment, training & competence as well as for barriers. The risk profile within technique & equipment demonstrated that most of the risks prior to the implementation have been dealt with. Training of the personal drastically decreased the risk profile within training & competence. However, increased risks were found for the source barriers as a consequence of the introduction of new barriers, which reduced the accessibility of laboratory results. This was a consequence of the Swedish legislation for patient related information in medical records, which prevents full access of a patient's laboratory test data without the patient consent.

The risk profile for the source; rules/policies/procedures is unchanged. The numbers of risks have decreased but the risks with a high score, which were identified prior to the implementation, were never dealt with.

Conclusion

Risk analysis is a useful method for identifying and managing possible patient safety risks and also a method for proactive work in health care organisations. Systematic and regular risk analysis is preventive of managing new routines. This may be a constructive tool for the patient safety work

References

[1] Stalhandske E, DeRosier J, Patail B, Gosbee J. How to make the most of failure mode and effect analysis, Biomed Instrum Technol. 2003 Mar-Apr;37(2):96-102

Address for correspondence Patient safety office in county council of Ostergotland
University hospital, S-581 85 Linkoping, Sweden
Annica Öhrn e-post: annica.ohrn@lio.se

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Multi-label Text Classification of German Language Medical Documents

Stephan Spat[a], Bruno Cadonna[a], Ivo Rakovac[a], Christian Gütl[b], Hubert Leitner[c], Günther Stark[c], Peter Beck[a]

[a] *Institute of Medical Technologies and Health Management, JOANNEUM RESEARCH Forschungsgesellschaft mbH, Graz, Austria*
[b] *Institute for Information Systems and Computer Media, Graz University of Technology, Graz, Austria*
[c] *Steiermärkische Krankenanstaltenges. m.b.H., Graz, Austria*

Abstract and objective

At nearly every patient visit, medical documents are produced and stored in a medical record, often in an unstructured form as free text. The growing amount of stored documents increases the need for effective and timely retrieval of information. We developed a multi-label text classification system to categorize free text medical documents (e.g. discharge letters, clinical findings, reports) written in German into predefined classes. A random sample of 1,500 free text medical documents was retrieved from a general hospital information system and was manually assigned to 1 to 8 categories by a domain expert. This sample was used to train and evaluate the performance of 4 classification schemes: Naïve Bayes, k-NN, SVM, and J48. Additional tests of the effect of text preprocessing were done. In our study, preprocessing improved the performance, and best results were obtained by J48 classification.

Keywords:

machine learning, classification, medical records, multi-label

Introduction

At nearly every patient contact with healthcare-providers, medical documentation is generated and stored in medical or nursing records, often as free text. With the increasing amount of stored, unstructured free text information, the need for effective and timely retrieval of relevant information is growing. In this work, we describe the development and the evaluation of an information system for multi-label classification of medical documents into predefined classes.

Methods

A random sample of 1,500 unstructured, free text documents written in German was extracted from an electronic medical record (EMR) of a general hospital in Austria. A domain expert (physician) manually classified the retrieved documents into one or more of the following classes: surgery, vascular surgery, casualty surgery, internal medicine, neurology, anesthesia and intensive care, radiology and physiotherapy. In average, 1.47 labels were assigned to a document. We built an automated multi-label text classification system in Java based on Weka [1], an open-source machine-learning framework. Four different kinds of classification schemes were compared: Naïve Bayes, k-NN, SVM and J48. 10-fold cross validation was used for evaluation. Moreover, the influence of text preprocessing (e.g. stop-word-removal, stemming, lowercasing) was studied.

Results and conclusion

Results for the F-measure [2] with and without preprocessing are shown in figure 1. J48 performed best, followed by 1-NN, SVM, and Naïve Bayes. The results were improved by text preprocessing. The best classification scheme (J48) with text preprocessing achieved an F-measure of 0.886.

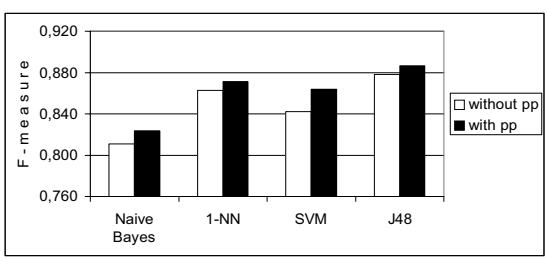


Figure 1 - F-measure with and without text preprocessing (pp)

Results show that it is possible to classify medical documents written in German originating from a general hospital with automated machine-learning classification schemes with promising results, comparable with [3]. This classification system is used in a prototype of an information retrieval system for score-calculation, thus influencing the display order of search results. Further studies are needed to evaluate the accuracy of the developed system in other hospitals as well as the user-perceived benefits of this prototype.

References

[1] Witten IH, Frank E: Data Mining: Practical machine-learning tools and techniques, 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

[2] Hripcsak G, Rothschild AS: Agreement, the F-Measure, and Reliability in Information Retrieval. J Am Med Inform Assoc. 2005; 12(3): 296-298.

[3] Wilcox A, Hripcsak G: Classification algorithms applied to narrative reports. Proc AMIA Symp. 1999; 455-9.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

What Health Influences are Caused by EMR Working?- In Case of Japanese Nursing Situation

Yukie Majima [a], Yasuko Maekawa[a]

[a] *School of Nursing, Osaka Prefecture University, Osaka, Japan*

Abstract

The purpose of this study was to determine how nurses used visual display terminals (VDT) and whether the use of VDTs had any influence on their health. In our study, the use of the EMR and the health influences of introduction of the EMR were surveyed using a questionnaire in a general hospital. The results indicated that VDT works were conducted in nurse stations rather than at the bedside. Only about 10% of nurses did exercises to prevent neck, shoulder and back pain before starting work. Since more than half of the nurses remained standing while inputting, extra guidelines should be developed for this aspect as the MHLW guidelines currently only deal with seated operations. Health influences after introduction of the EMR, more than half the subjects complained of worsened symptoms including eye strain, stiff shoulders and neck, and general fatigue, all of which are signs of VDT syndrome.

Keyword:

nurses, VDT works, electronic medical record system, health influences

Introduction

Currently in Japan, the use of EMR and digital ordering systems in hospitals is increasing. Since medical service is one of the seven main fields of the e-Japan priority policy program, it is assumed that digitalization will be promoted in medical facilities. Accordingly, appropriate workplace health measures should be taken in response to changes in the working environment of nursing personnel.

Methods

1) Study period: March 8 to 29, 2005. 2) Subjects: Nursing care personnel (378 individuals) working at a private general hospital (477 beds) in Japan which had introduced an EMR around a year previously. 3) Method and contents of study: Self-reported questionnaire study. A questionnaire was delivered to the subjects and was recovered on completion under ethical cares. It consisted of questions regarding the VDT working situation of EMR((1)posture, (2)place, (3)adjust materials around PC, (4)input device, (5) time per one use of EMR, (6)brightness around PC on each shift, (7)setting-up exercise before working) and the health influences (eye strain, low vision, dried eyes, stiff neck and shoulders , numbness of fingers, back pain, fatigue, headache, stress, gloomy) after using the EMR.

Results

323 questionnaires delivered were recovered (85.4% recovery). 193 subjects (59.8%) were in their 20s, 80 (24.8%) were in their 30s, 27 (8.4%) were in their 40s and 19 (5.9%) were in their 50s. Four subjects (1.2%) did not reveal their age. The average number of years working at the hospital was 5.67.

1) The VDT working situation of EMR

VDT works were conducted in nurse stations rather than at the bedside. Since more than half of the nurses remained standing while inputting records.

2) The health influences after introduction of the EMR

Only about 10% of nurses did exercises to prevent neck, shoulder and back pain before starting work. More than half the subjects complained of worsened symptoms including eye strain, stiff shoulders and neck, and fatigue, all of which are signs of VDT syndrome (Figure 1).

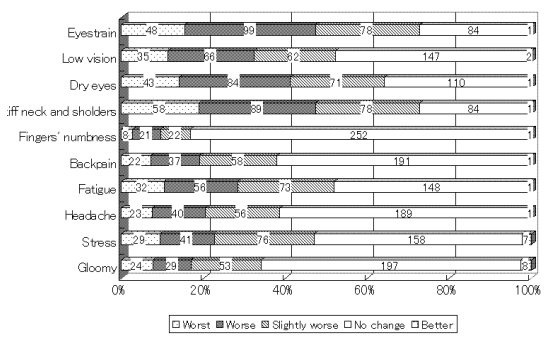


Figure 1 - Subjective signs of health influences after using EMR

Conclusion

Nurses' VDT working situations are different form those of office workers etc. Workplace health measures for VDT operations by nurses should be urgently revised based on the results of this study.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Using PDA to Transform the Long MDS-HC Evaluation Form into a Favored System

Chiao-Ling Hsu, RN, MS, Yu-Yin Kuo, RN, MS, Polun Chang, PhD

Institute of Biomedical Informatics (formerly Health Informatics and Decision Making), National Yang-Ming University, Taiwan

Abstract

The MDS-HC has been an effective home care evaluation form. However it was not yet accepted in Taiwan because it is too long for our over-burdened home care nurses. We used a self-developed PDA information representation model to design the PDA MDS-HC support system and used the Technology Acceptance Model to examine its potential acceptability. The results showed a well-designed PDA could greatly improve the usability of a originally un-favored paper system.

Keywords:

MDS, home care, PDA, information representation

Introduction

The Resident Assessment Instrument, consisting of a Minimum Data Set (MDS) evaluation form, Client Assessment Protocols (CAPs) and Triggers, has been well internationally known, applied, and even used for the insurance claim, in the long term care [1]. The MDS-HC is vary complete and covers 17 main categories and as many as 64 evaluation items. However, currently its professional completeness doesn't make it a formal evaluation form for the home care in Taiwan because it is too long for our over-burdened home care nurses.

PDA has been a potentially useful tool for mobile nursing, especially for those in home care settings. But its interface features, such as difficult data entry in writing and small display screen, still confine its use. We developed a PDA information representation approach which had been practically proven to be easy to use by nurses [2][3]. A modified version was latter developed for the long forms.

The purpose of this study was to use the modified PDA information representation approach to design the PDA-based MDS home care evaluation support system and to examine the home care nurses' acceptance of the PDA-based MDS-HC system.

Methods

The modified PDA information representation approach is mainly composed of 4 principles:

1. The screen display area is separated into three sections for the users to clearly see and to easily navigate through the form.

2. User should be able to switch to any item in less than 3 taps.

3. The hand-writing should be replaced, if possible, by tapping for data entry.

4. User should be well reminded what tasks have been done.

A team of nursing user, nursing programmer and experienced PDA programmers was organized and the prototyping approach was used to develop the system. A convenient sample of 24 subjects, who were trained of the MDS, was used for testing. 3 representative home care scenarios were written for users to test the PDA system. Davis' Technology Acceptance Model was used to evaluate users' acceptance of the system in terms of their perceived ease of use and usefulness.

Results

The representative PDA screen was shown in Figure 1, in which uses could navigate the system through tapping the upper main category and subcategory buttons, and enter data in the lower half data entry area. The highlighted buttons mean answered questions.

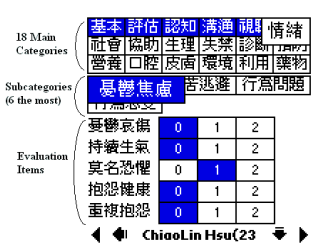


Figure 1. The representative screen shot.

Figure 1 - The representative screen shot

58% of subjects never used the MDS-HC form. Most of the subjects have BS degree and aged from 20-50. 100% agreed the system was easy to use and 92% agreed the system was useful.

Conclusions

Our study showed that a well-usability-engineered PDA system can well improve the usability of both the professional forms, which was too long to be practically accepted, and PDA, which small-display interface usually hinders its popularization.

References

[1] Morris JN, Fries BE, Bernabei R, Steel RK, Ikegami N, Carpenter GI. RAI-Home Care (RAI-HC) assessment manual for version 2.0. Boston: Inter RAI; 1999.

[2] Chang P, Hsu YS, Tzeng YM, et al. Development and Pilot Evaluation of User Acceptance of Advanced Mass-Gathering Emergency Medical Services PDA Support Systems. MedInfo 2004 Symposium, p:1421-1425.

[3] Chang P. Tzeng YM, Wu SC, et al. Development and Comparison of User Acceptance of Advanced Comprehensive Triage PDA Support System with a Traditional Terminal Alternative System. American Medical Informatics Association 2003 Symposium, p:140-144.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

A Sustainable, Multi-Organizational Model for Decision Support During Public Health Emergencies

Julie J. McGowan[a,b], Margaret W. Richwine[a], J. Marc Overhage[a,b,c]

[a] *School of Medicine, Indiana University, Indiana*
[b] *Regenstrief Institute, Inc., Indiana*
[c] *Indiana Health Information Exchange*

Abstract

In an effort to provide decision support during times of public health emergencies, Regenstrief Institute, Inc. and the Indiana University School of Medicine, in partnership with the county and state health departments, created a sustainable model to deliver information to health care providers and public health workers. The model leverages extant systems and processes, including active surveillance through electronic laboratory reporting, delivery of health information as part of the Indiana Health Information Exchange, and evidence based utilities and Blog technology to create a public health utility with disease specific information and epidemiologic reporting requirements.

Keywords:

disease outbreaks; bioterrorism; information dissemination; decision support systems, clinical

Introduction

Knowledge at the point of decision making is critical to improving health care and this is especially true in public health when medical decisions in response to a bioterrorist event or an emerging epidemic could have a much greater impact. However, the ability to access such information has been extremely problematic because of the different levels of connectivity and training available to public health workers and health care providers across large areas.

To address this issue a partnership was formed among the Indiana State Department of Health and the Marion County Health Department, the Indiana University School of Medicine Libraries (IUSML) and the Indiana Health Information Exchange to establish a cost-effective model for both alerting health care providers and public health workers to developing threats and providing them the needed information for both treatment and reporting.

Methods

The model, predicated on an earlier construct [1], was built on three existing programs. The Regenstrief Institute, Inc. in conjunction with the Indiana Network for Patient Care and the Indiana State Department of Health developed an active surveillance of reportable conditions program with electronic laboratory reporting. [2] I3, the Indiana IAIMS (Integrated Advanced Information Management Systems) Initiative, through the Indiana Health Information Exchange and its DOCS4DOCS utility, provided a framework to notify health care providers about real or suspected public health emergencies. [3] The IUSML offered evidence based decision support using Web technologies.

The goals of the project were to insure rapid notification of public health problems to a large yet appropriate group of health care providers and to provide access to quality filtered knowledge supporting treatment guidelines and reporting requirements. Because the eventual objective was to operationalize the model for long term adoption, it was essential to develop sustainable methods.

Electronic laboratory reporting uses LOINC coding and HL7 messaging to provide timely information to the health departments regarding the potential for a public health event. Notifiable conditions from laboratories around the Indianapolis metropolitan statistical area are reported to the health departments. The enhancement model facilitates provider notification and the creation of decision support information.

The health care provider indexing in the Indiana Health Information Exchange enables the rapid notification of health care providers about emerging public health threats based on location and/or condition. This insures a higher level of relevance of the warning for the provider. Simultaneously, the IUSML is contacted about the suspected problem and receives information concerning proposed optimum treatment and management guidelines and the required reporting of the specified condition.

The IUSML is responsible for searching for current evidence about management of the condition and creating a web presence containing this information in a user effectiveness format. In addition, reporting requirements and links to the health department web sites are included. To insure sustainability, a Blog utility was chosen because of its ease of document creation and maintenance, the famil-

iarity of the format for most users, and its functionality in searching for prior conditions.

In addition to the reporting and management information for health care providers, a listserv is maintained for notification of public health workers. The health departments have their own systems for notifying their affiliated sites, however the notification of access to knowledge-based information is handled by the IUSML once the Blog site has been updated and approved.

While most of the information about emerging public health events comes from the electronic surveillance functionality, some comes from the public press. While this usually involves very specific localities and has little potential to develop, decision support is also beneficial for these events and the Web site is updated as the need presents itself.

Results

As of November 2006, there were thirteen public health events that prompted the creation of knowledge support on the Web site. The mechanism for the notification of health care providers has been developed using extant resources but has as yet to gain widespread adoption, not because of the technology but because of organizational issues.

The process for the notification of public health workers was implemented with the second public health event posted and, as a result, the utility has achieved wide spread use. Anecdotal responses to the initiative have indicated that the Web site has provided critical information for both management of health care conditions and epidemiologic reporting.

Discussion

By using three extant processes, active surveillance through electronic laboratory reporting, DOCS4DOCS delivery of information to health care providers, and the evidence based medicine services of the IUSML, a sustainable model was developed to deliver decision support information to targeted health care providers and public health workers during instances of public health events.

The technology and processes were easily adapted to meet the needs of the health departments and system users. However, organizational issues regarding the health care provider notification component have precluded the full operationalization of the system. It is anticipated that these issues will be resolved within the next year and that the system will be fully implemented. However, the public health worker notification and the creation and use of the Blog utility have proven to be an effective means to providing decision support during times of emerging infectious diseases and bioterrorism events.

Conclusion

Emerging public health crises require coordination of information from a variety of sources and targeted provision of quality filtered knowledge to a wide range of health care workers. Access to evidence is critical to the management of the events for both the population and individual patients; collecting information to monitor the impact of the occurrences contributes to the knowledge base and helps mitigate future occurances. Using extant technology and promoting organizational partnerships can offer a sustainable model to enhance responsiveness to these events.

Acknowledgements

The authors performed this work at the Regenstrief Institute for Health Care and the Indiana University School of Medicine and they were supported in part by the National Library of Medicine, grant number GO8 LM008232-02.

References

[1] Kay BA, Timperi RJ, Morse SS, Forslund D, McGowan JJ, and O'Brien T. Innovative information-sharing strategies. Emerg Infect Dis 1998: 4(3): 465-6.

[2] Overhage JM, Suico J, and McDonald CJ. Electronic laboratory reporting: barriers, solutions and findings. J Public Health Manag Pract 2001: 7(6): 60-6.

[3] McGowan JJ, Overhage JM, Barnes M, and McDonald CJ. Indianapolis I3: the third generation Integrated Advanced Information Management Systems. J Med Lib Assoc 2004: 92(2): 179-87.

Address for Correspondence

Julie J. McGowan, Ph.D., FACMI
IU School of Medicine, IB-310
975 W. Walnut St.
Indianapolis, IN 46202-5121 USA

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Customized Early Warning System Based on HTN for Home Healthcare Model

Seung-Jin Jang[a], Jip-Min Jung[a], Sung-Oh Hwang[b], Young-Ro Yoon[a]

[a] *Department of Biomedical Engineering, Health Science College, Yonsei University, Wonju, South Korea*[b]*Department of Emergency Medicine, Wonju College of Medicine, Yonsei University, Wonju, South Korea*

Abstract and objective

We adopted hierarchical task network (HTN) planning in a customized early warning system for home healthcare model. It is necessary to design a customized early warning system for the patients who were in various health states, because regular or irregular report format and alarm delivery systems were diversified according to severity of health state. HTN planning is suitable for use of constraint programming so as to effectively prune the search space during the search for solution. An efficient and scalable information control is presented by use of HTN. The paper also briefly deals with a process strategy for the early warning system.

Keywords:

hierarchical task network, emergency response, home healthcare model

Introduction

As the elderly population and the demand of well-being life increasingly grow, IT-based technology allowing bio-signal measurement and assessment at home have been focused nowadays. The elderly or citizen can be efficiently controlled with respect to health status parameters which act as input features in the early warning system for health status.

Methods & results

The early warning system developed by Home Healthcare Management System Research Center (H2MSRC) in Yonsei University adopted HTN planning in order to efficiently design the planning of emergency response because of extendible, intelligible, and easily communicated properties. Events derived from 7 objective status parameters measured from 4 types of sensor devices and 3 subjective parameters results in customized emergency response according to the formalized SHOP domain model. (Ref. figure 1)

Conclusion

HTN planning was practically adopted in our early warning system for Home healthcare model and revealed the intelligent and collaborative planning ability.

Acknowledgment

This study was supported by a grant of the Korea Health 21 R&D Project, Ministry of Health & Welfare, and Republic of Korea (02-PJ3-PG6-EV01-0001).

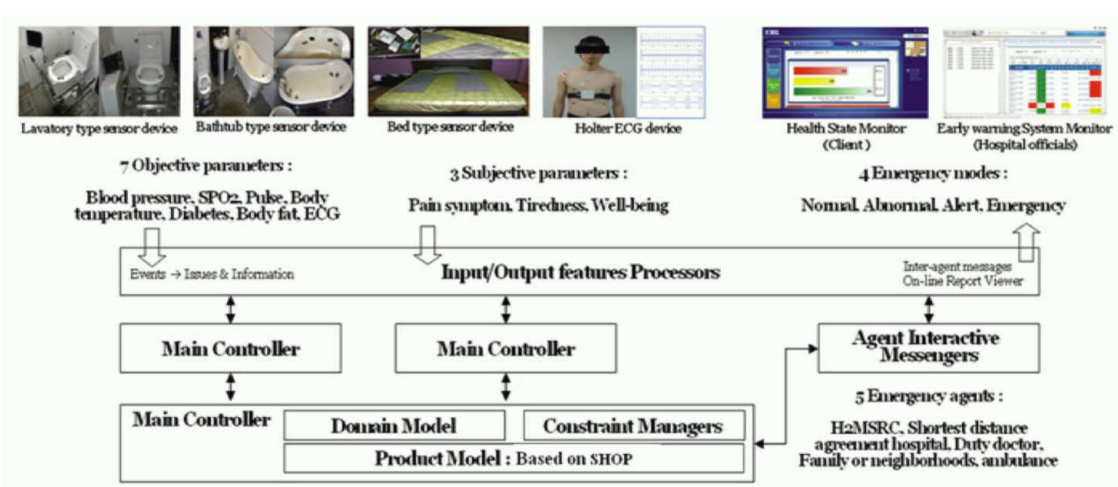


Figure 1

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Open Source Patient Data Management System for Intensive Care

Massaut J, Reper P, Hooghe L, Gottignies P

Surgical Intensive Care, Brugmann Hospital, Université Libre de Bruxelles

Abstract and objective

In Intensive Care Units, the amount of data to be processed for patients care, the turn over of the patients, the necessity for reliability and for review processes indicate the use of Patient Data Management System (PDMS). To respond to the needs of a Surgical Intensive Care Unit, we developed a PDMS based on open source software and components.

The software was designed as a client–server architecture running on the Linux operating system and powered by the PostgreSQl data base system. The client software was developed in C. The application offers the following functions: medical notes captures, observations and treatments, nursing charts with administration of medications and scoring systems functionalities. The PDMS was used to care more than two thousands patients with the expected reliability and functionalities.

Key words:

database management system, software, intensive care

Introduction

Patient Data Management Systems are mandatory in Intensive Care Unit in response to the amount of data to be processed, the turn over of the patients and the necessity for review processes. To respond to the needs of our unit we developed a PDMS based on open source software and components.

Methods

The software was designed as a client-server architecture running on the Linux operating system (SUSE Linux Enterprise Server 8.0), using the PostgreSQL relational database (v 7.2). The client software was developed in C using the GTK interface library. Remote access from remote PCs is implemented by virtual network connections (VNC), the use of VNC servers and VNC viewers.

The hardware consists in two Intel x86 servers (one master and one slave to assure the integrity of the database by replication),

14 medical grade panel PCs connected via RS232 medical bus to the patient's monitoring devices and the servers via local Ethernet network.

The software, developed on the Linux platform, offers the following functions: medical notes captures with patient's history, observations and treatments, nursing charts with functionalities for administration of medications, and scoring system possibilities for patient's classification. Interoperability between these modules is realized through access to the PostgreSQL database and not the use of local memory in the interface. The software was developed to be open source in all its components.

Results

The PDMS is in used in our unit from February 2004 and was used to care more than two thousands patients. The system is accessible at every bed through panel PCs and at desks or offices through VNC viewers on windows PCs. Its design allows an access to the database's functionalities with a high availability level (less than 5 hours of interruption over one year).

The use of open source resources was effective to customize the solution to ICU's request. The use of the C language permitted to obtain small response times but limits the portability of the system and complicated the debugging process in this critical environment.

Conclusion

PDMS based on open source software components are effective and able to respond to the needs of the ICU environment, with a high availability level.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Why Teach Computer Security to Medical Students?

Ana M. Ferreira [ac], Ricardo Cruz-Correia [ab], Altamiro Costa-Pereira [ab]

[a] *CINTESIS - Center for Research in Health Information Systems and Technologies, Faculty of Medicine of Porto* [b] *Biostatistics and Medical Informatics Department, Faculty of Medicine, University of Porto* [c] *Department of Informatics, Faculty of Medicine, University of Porto*

Abstract

The introduction of Electronic Medical Records (EMR) within the healthcare practice can be beneficial in order to integrate and centralize heterogeneous patient information. However, there are still some problems that hinder the successful use of EMR. The concern for patient privacy is one of them. The aim of this paper is to present the results of a study that assesses attitudes of 1st year medical students towards computer security and the EMR. An anonym questionnaire was given to the students at the beginning and at the end of the academic year of 2003/2004 for them to comment on several security related scenarios. 238 questionnaires were answered at the beginning of the year whilst 222 were answered at the end of the year. The students feel, at the end of the year, that they understand better what computer security is and how to protect patient privacy information. This shows that teaching computer security to medical students, the future users of EMR, can greatly influence the success of EMR integration and therefore improve and fasten healthcare treatment.

Keywords:

computer security; education, medical, undergraduate

Introduction

The introduction of Electronic Medical Records (EMR) within the healthcare practice allows for the integration of heterogeneous information that are usually scattered over different locations [1] [2]. However, there are some barriers that impede its successful integration in most healthcare practices [3] [4].

One specific barrier relates with the privacy and security of patient information [5]. The use of new information systems within healthcare stresses the need for young doctors to comprehend them from their conception so that they can be used in a beneficial way and support their future daily work. As such, all the feedback provided during their training into the medical profession is essential for the enhancement of those systems [6], moreover in terms of computer security.

The Biostatistics and Medical Informatics Department of Porto Faculty of Medicine teaches Ethics and Medical Informatics to 1st year medical students [7]. The later subject includes theoretical and practical lectures about Electronic Medical Records (EMR) and computer security.

This study aims to assess the opinions, attitudes and awareness of 1st year medical students towards computer security issues relating to EMR, and how these can affect the successful integration of EMR within the healthcare practice.

Methods

An anonym questionnaire was given to the students both at the beginning and again at the end of the academic year of 2003/2004. It was applied two times so that we could compare their attitudes before and after they had attended the Ethics and Medical Informatics' subjects.

The questionnaires introduced 3 scenarios for the students to comment. The first scenario described a breach of patient privacy to an EMR by one of their colleagues. There were two questions relating with this scenario:

- Q1.A – Is the described scenario a security breach?
- Q1.B – What would you do if you found out about this breach?

The second scenario included additional information to the first scenario. It explained that the colleague in question had shared his password with a friend and that friend was the one to access patient private information, without him knowing it. The question related with this scenario was:

- Q2 – Would you change your attitude if you found out this new piece of information?

The third scenario introduced the issue of more sensitive information (e.g. HIV, Cancer results or even VIP related) and how this information must be protected. The question presented within the questionnaire was the following:

- Q3 – Do you think this kind of information requires stronger security measures than other types of information?

The answers to these questions were inserted into SPSS® and analysed separately.

Results

A total of 460 questionnaires were filled by the students. 52% (238) were answered before the lectures started whilst 48% (222) after the lectures finished. Table 1 shows the results obtained from the applied questionnaires.

Table 1 – Results obtained from the questionnaires

		Before the lectures % (n)	**After the lectures % (n)**
Answered questionnaires		*(238)*	*(222)*
Q1.A	*Valid answers*	*98 (232)*	*98 (217)*
	Yes	100 (232)	100 (217)
Q1.B	*Valid answers*	*61 (144)*	*60 (132)*
	Reason with	54 (77)	44 (58)
	Inform	40 (58)	50 (66)
	Others	6 (9)	6 (8)
Q2	*Valid answers*	*62 (148)*	*62 (138)*
	No	74 (109)	83 (115)
	Yes	26 (39)	17 (23)
Q3	*Valid answers*	*43 (103)*	*91 (204)*
	No	44 (91)	38 (77)
	Yes	55 (112)	62 (127)

For Q3, the main reason given by the students that felt no extra security measures were needed to access more sensitive information is that all security measures must be effective for all cases, independently of the patient or healthcare performed. The majority of the students that thought extra security measures were necessary agreed that this would provide for the protection of certain social groups from discrimination.

Conclusion

According to this study's results, after Medical Informatics and Ethics' lectures, students feel more conscientious to report privacy breaches to responsible parties (Q1.B). They understand better what computer security is and how to behave in order to protect confidentiality of electronic information. They consider indirect disclosure of sensitive information, such as with another person's password, a serious fault (Q2). Further, at the end of the year, students become more aware for the need of different protection levels of security depending on how sensitive information can be (Q3).

We believe that the introduction of Medical Informatics and Ethics early in the degree of the Medical course has an influence in the awareness and attitudes of first year medical students towards computer security and EMR. This can greatly influence the success of EMR integration whilst improving and fastening healthcare treatment.

References

[1] Waegemann C. EHR vs. CPR vs. EMR. Healthcare Informatics online. 2003.

[2] Cruz-Correia R, Vieira-Marques P, Costa P, Ferreira A, Oliveira-Palhares E, Araújo F, et al. Integration of Hospital data using Agent Technologies – a case study. AICommunications special issue of ECAI. 2005;18(3):191-200.

[3] Sprague L. Electronic health records: How close? How far to go? NHPF Issue Brief. 2004 Sep 29(800):1-17.

[4] Miller RH, Sim I. Physicians' use of electronic medical records: barriers and solutions. Health Aff (Millwood). 2004 Mar-Apr;23(2):116-26.

[5] Knitz M. HIPPA compliance and electronic medical records: are both possible? Graduate research report: Bowie State University. Maryland in Europe; 2005.

[6] Davis L, Domm J, Konikoff M, Miller R. Attitudes of First-year Medical Students Toward the Confidentiality of Computerized Patient Records. JAMIA. 1999; 6: 53-60.

[7] Freitas J, Cruz-Correia R, Almeida F, Costa-Pereira A. Evolution of Medical Informatics teaching in a medical undergraduate course. MedNet 2003, Eighth World Congress on the Internet in Medicine; 2003; Geneva - Switzerland; 2003. p. 312-4.

Author Correspondence

Ana M. Ferreira - First author's contact: CINTESIS – Center for Research in Health Information Systems and Technologies, Faculty of Medicine of Porto, Al. Prof. Hernani Monteiro, 4200-319 Porto, Portugal
(Phone: +351 22 551 3613; Fax: +351 22 551 3613; email: amlaf@med.up.pt)

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Application of Wireless and Mobile Computing Technologies to Improve the Efficiency of Patient Care and Education: The Role of Medical Engineering and Information Technology

Lin Guo

Centre for Adult and Paediatric Gastroenterology, Institute of Cell and Molecular Science, Queen Mary's School of Medicine and Dentistry Barts and The London, , Queen Mary, University of London, UK.

Abstract

This study explored the potential of the application of wireless and mobile computing technologies to be used in improving the efficiency of patient care and education and future developments in information and communication technologies to support healthcare professionals and medical students in healthcare research, medical education and training. The design used for this study was a systematic review of published materials obtained from EMBASE and MEDLINE online databases, and the Cochrane Library database, including personal observations. Today, more than 50% of healthcare professionals and medical students are using Personal Digital Assistant with expected growth of more than 75% by year-end 2007. In addition, wireless and mobile computing technologies allows Personal Digital Assistant to connect directly to networks or the Internet. Studies relating to processes of patient care and should evaluate mobile computing technologies as a potential timesaving tool. Wireless and mobile computing technologies is only beginning to take its first step in improving patient care and education. They have shown a positive impact on patient safety, health care efficiency, and ultimately patient satisfaction.

Keywords:

wireless and mobile computing technologies, patient care and education, efficiency, personal digital assistant

Introduction

Wireless is a term used to describe telecommunications in which electromagnetic waves (rather than some form of wire) carry the signal over part or all of the communication path.[1] Mobile Computing is a generic term describing your ability to use technology 'untethered', that is not physically connected, or in remote or mobile (non static) environments.[2] The term of wireless and mobile computing technologies is evolved in modern usage such that it requires that the mobile computing activity be connected wirelessly to/through the application of internet or to/through a private network. This connection ties the mobile device to centrally located information and/or application software through the use of battery powered, portable, and wireless computing and communication devices.[1, 2] This includes devices like laptops with wireless LAN or wireless WAN technology, smart mobile phones, wearable computers and Personal Digital Assistants (PDAs) with Bluetooth or IRDA interfaces.

This study explored the potential of the application of wireless and mobile computing technologies to be used in improving the efficiency of patient care and education and future developments in information and communication technologies to support healthcare professionals and medical students in healthcare research, medical education and training.

Design and methods

The design used for this study was a systematic review of published materials obtained from EMBASE and MEDLINE online databases, and the *Cochrane Library* database, including personal observations. Materials that match the set criteria on the application of wireless and mobile computing technologies in medical engineering and information technology research for healthcare professionals and medical students to improve the efficiency of patient care and education were selected and analysed following the United Kingdom National Health Service Centre for Reviews and Dissemination Guidelines. A variety of data collection approaches were developed to ensure data were collected on the various aspects.

Results

Wireless and mobile computing technologies is seen to be convenient to get in touch, compact, fast and portable, but problems are attached to the levels of security, confidentiality and scalability of the hardware. It is also apparent that most commonly used wireless and mobile computing technologies within the health contexts is the Personal Digital Assistant. Today, more than 50% of healthcare professionals and medical students are using Personal Digital Assistant with expected growth of more than 75% by year-end 2007.[3] Not only can the healthcare professionals and medical students use the tool on the Personal Digital Assistant as a quick look-up resource at the bedside, but can also flag a topic for further research when back in the office. Changes in treatment guidelines, concerns about patient safety, efforts to contain costs, time limitations, and better informed patients make it critical to have clinical reference information at the point of patient care and edu-

cation.[4] In addition, wireless and mobile computing technologies allows Personal Digital Assistant to connect directly to networks or the Internet.

Discussion

According to the results of this systematic review of this study, wireless and mobile computing technologies has become a valuable resource for both healthcare professionals and medical students over the past decade. This has enabled new perspectives to be developed on the interface between hardware, software and education and training processes for those involved in delivering healthcare research. Studies relating to processes of patient care and should evaluate wireless and mobile computing technologies as a potential timesaving tool, as they can be synchronised with hospital information systems to facilitate retrieval of patient information. The findings also indicate that the healthcare professionals and medical students would benefit from some technologies that they can easily use to search for such sources as database, E-journals and the Internet. At a theoretical level, wireless and mobile computing technologies such as Personal Digital Assistant is ideal for meeting these needs. Further study on processes of patient care and education should also explore wireless and mobile computing technologies as vehicles for disseminating evidence-based guideline recommendations.

Conclusion

Wireless and mobile computing technologies is only beginning to take its first step in improving patient care and education. They have shown a positive impact on patient safety, health care efficiency, and ultimately patient satisfaction. The integration of the Internet and wireless and mobile solutions will transform the use of information and communication technologies in patient care and education and take it in the role of medical engineering and information technology. The potential of wireless and mobile computing technologies is vast and its principles, application and practices are seen as a choice for healthcare professionals and medical students to be in the right place at the right time. In the future we will see more valuable resources in improving patient care and by developing wireless and mobile computing technologies.

References

[1] https://SearchMobileComputing.com, accessed on 4th November 2006

[2] http://en.wikipedia.org/wiki/Mobile_computing, accessed on 5th November 2006

[3] Osonnaya C and Osonnaya K. Information management in healthcare: analysis of trends and issues. International Journal of Medicine 2004; 1: 5-8

[4] Guo L and Osonnaya C. Improving the efficiency of patient care and education: the role of medical engineering and technology in gastroenterology. International Journal of Medicine 2006; 3: 149-151. (In press)

Address for correspondence

Mr Lin Guo
Centre for Adult and Paediatric Gastroenterology
Institute of Cell and Molecular Science
Queen Mary's School of Medicine and Dentistry
Barts and The London
Queen Mary, Uiversity of London
Whitechapel, London
E1 2AL
UK
E-mail: lin.guo@qmil.ac.uk or guolinlondon@yahoo.co.uk

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Comparing Messages in an Online Communication Forum for Cancer Patients with Patients' Messages to a Clinical Nurse Specialist

Annette Jeneson MA, Trine Andersen RN, Cornelia Ruland RN, Ph.D

Center for Shared Decision Making and Nursing Research, Rikshospitalet-Radiumhospitalet Medical Center, Oslo, Norway

Abstract

WebChoice is an online support system where cancer patients can exchange messages with other patients in an open communication forum as well as send personal e-mails to a clinical nurse specialist (CNS) who responds to their concerns. We compared the content of messages posted in the open forum with personal e-mail messages sent to the nurse. While patients were concerned with similar topics in both communication areas, there were differences in the types of messages sent. More patients actively used the patient-nurse email service compared to the forum, suggesting that nurses can play an important role in online systems supporting cancer patients over the Internet.

Keywords:

online systems, communication, nurse-patient relations, neoplasms

Introduction

Despite a growing interest in online support forums and patient-provider communication systems, knowledge of the relative use and benefits of these communication systems is still limited. In particular, there is little information about how patients' communications with care providers online differ from communications in discussion groups. Additionally, while a few studies have analyzed online messages between patients and physicians, online patient-nurse communication and the potential for nurses to support patients via the Internet is as of yet largely unexplored. The purpose of this study, therefore, was to compare the content of patient-nurse messages to messages posted on an online discussion forum for cancer patients. The study is part of a larger ongoing trial testing the effects of WebChoice, an internet support system for breast- and prostate cancer patients [1]. WebChoice includes tools for symptom monitoring, tailored information to support self-management, support forums for anonymous group discussion, as well as a more "private" communication area to exchange messages with a CNS specializing in cancer care.

Methods

We examined forum postings and e-mail messages by patients who logged on to WebChoice at least once from March - October 2006, resulting in a sample of 355 postings and 174 e-mails. Messages were coded according to the 'type' and 'topic' categories listed in Table 1 below. The 'type' categories are consistent with the coding schema of Klemm *et al* [2]. A message could belong to more than one category. An e-mail with questions about lymphedema and anxiety, for example, would be coded as *information seeking* for 'type', and *symptoms* and *feelings* for 'topic'. Messages were independently coded by two of the authors (AJ and TA) and 10% were coded by both to compare interrater agreement, established at 98 % for e-mail messages and 97 % for forum postings.

Results

Table 1 shows the percentages of e-mails and forum postings coded under the various 'type' and 'topic' categories.

Table 1. Percentage of messages categorized by type and topic

		E-mails (*174 total*)	**Postings** (*355 total*)
TYPE	Personal experiences/ opinions	66	75
	Information giving	2	20
	Information seeking	62	20
	Encouragement / Support	0	17
	Thanks	22	8
TOPIC	Health personnel / institutions	21	20
	Treatment / Tests / Test results	63	59
	Symptoms / Side-effects	64	66
	Energy / Fatigue / Sleep	9	10
	Feelings	14	10
	Sexuality / Partner	3	7
	Family / Colleagues / Others	5	3
	Living with cancer / Lifestyle	22	23
	Metastasis / Relapse	7	8

The average number of topic categories per e-mail or posting was 2.4 (±1.5) and 2.3 (±1.3), respectively. As seen in Table 1, patients were concerned with largely the same topics in their e-mail messages and in their forum postings, but the type of message or posting varied, with e-mails being more 'information seeking' in nature and postings being more 'supportive' and 'information giving' to others.

There were also differences in use of the e-mail component of the WebChoice system compared to use of the discussion forum in terms of the number of participants actively using these two different components. 71% of participants given access to WebChoice logged on at least once in the data collection period (43 females and 31 males). Of these active WebChoice users, more patients used the patient-nurse communication area (n=45 or 61% of WebChoice users) compared to those submitting messages to the forum (n=34 or 46% of users). This holds true for breast cancer patients as well as prostate cancer patients. 72% of female users who logged on to the system at least once between March to October 2006 sent at least one e-mail to the nurse while 58% submitted at least one posting. 45% of active male users sent an e-mail to the nurse while 29% contributed to the forum. However, while more patients wrote e-mails to the nurse compared to the number of patients who submitted a posting in the forum, those patients actively participating in the discussion forum submitted on average more postings (10.4±10.5) compared to the average number of e-mails sent to the nurse (3.9 ± 4.6).

Discussion and conclusion

This sample of messages and forum postings, although small, provide useful information for health professionals interested in online communication systems. In a number of computer-based support systems for patients the communication areas are consistently highlighted as the most popular sections of the system, yet our knowledge about patients' use and benefits of different features within these communication areas is limited. A preliminary usage analysis of WebChoice suggest the forum is so far the most visited section in this support system [3], but the fact that more patients submit messages to the nurse via e-mail compared to the forum, and that the nature of messages in these two communication areas differ, suggest that opportunities for patient-nurse communication can provide valuable support for patients beyond that of support offered by participation in online discussion groups. Patients often experience multiple symptoms during treatment and rehabilitation, yet short hospital admissions allow little time for detecting and relieving these. Moreover, side-effects of certain treatments are often worst after the patients are discharged to home.

Therefore, patients could greatly benefit from support through an Internet based system, where they can communicate with a care provider independent of scheduled hospital or doctor appointments in an environment that is readily accessible and even anonymous. In this sample almost all of the questions patients asked via e-mail could be appropriately addressed by the CNS and did not require advice from a physician or other specialist. This suggests that online communication with a nurse may potentially reduce not only needless patient suffering and worrying but also the numbers of doctor's appointments scheduled. As cost concerns and shortages of health professionals continue to rise, online peer- and professional support provided by nurses could prove a viable health care supplement that can improve delivery of high quality patient care in the future [4].

Acknowledgments

We thank Laura Slaughter and Gro Hjelmeland Grimsbø for helpful discussion. The research was funded by the Norwegian "Health and Rehabilitation" Grant # 2003/3/0429.

References

[1] Ruland CM, Andersen R (2004) Designing WebChoice - Individualized Support for Cancer Patients through the Internet. Medinfo, San Francisco, Sept 5-11, 2004, p.1840.

[2] Klemm P, Hurst M, Dearholt SL, Trone SR (1999) Cyber Solace: Gender Differences on Internet Cancer Support Groups. Comput Nurs 17(2), pp. 65-72

[3] Ruland CM, Jeneson A, Andersen T, Slaughter L, Andersen R, Schjødt-Osmo B, Moore SM. Designing Tailored Internet Support to Assist Cancer Patients in Illness Management. Submitted for publication, AMIA Nov 10-14 2007, Chicago IL.

[4] Sorrells-Jones J, Tschirch P, Liong MAS (2006) Nursing and Telehealth: Opportunities for Nurse Leaders to Shape the Future. Nurse Leader 4(5), pp.42-58.

Address for correspondence

Professor Cornelia Ruland
Center for Shared Decision Making and Nursing Research
Rikshospitalet-Radiumhospitalet HF
N-0027 Oslo
Norway
E-mail: cornelia.ruland@rr-research.no
Telephone: (+47) 23 07 54 60

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Guideline-based Visualization of Medication in Chronic Disease

Ján Stanek, Michelle Joy Davy

Advanced Computing Research Centre, University of South Australia, Australia

Abstract

Chronic disease management at the General Practice level is a challenging task requiring synthesis of information across time and possibly several practitioners. Sparsity of data, lack of structure, lack of time are limiting factors of computer use in this domain. Authors are exploring the concept of visualization of individual patient's medication using a clinical guideline to provide some structure to the problem by creating a state-transition model. This approach was shown to be promising at the practice level in previous research [1,5,6] where an overall synthesis of practice decisions was created and alerts were generated on the basis of individual transactions or states - current focus is on individual patient and the sequence of states describing his/her medication history.

Keywords:

general practice, visualization, decision support, guideline

Introduction

In managing chronic diseases, the information on the past course of the disease can be important - but many practitioners see reviewing of past data as a (necessary) evil to be done as quickly as possible [2]. With more and more general practitioners (GPs) using a computer in their everyday practice [3] it is more than tempting to use the existing data to provide support for quality and continuity of care. Typically the data is meant to be used by humans so the data is rather unstructured, with highly variable quality in terms of completeness or adherence to some coding standards. In a comprehensive review it was shown, that prescription data is the one most complete and reliable [4]. Assuming, that medication reflects significant proportion of decision-making about a case a guideline-based state-transition model was created and used for analysis at the practice level, as well as for generation of patient-related alerts [1,5,6]. While this approach shows promising results, information on patients as drawn from the model does not take into account patient-specific sequences of states. In the current work we explore the visualization of these sequences as well as generating alerts based on sequential patterns for particular patients in line with the ideas of Aigner and Miksch [13].

Materials and methods

The backbone of the study was a state-transition model focusing on treatment of hypertension in diabetes mellitus based on an American Diabetes Association guideline [7]. The data used in the study were de-identified extracts from a general practice in rural Australia holding age, gender, visit dates, problem codes, blood pressure measurements and drug prescriptions. The extracts covered 6 years and contain more than 70000 prescriptions. Two sequences of events/states were generated for each patient: prescription path and treatment path. Prescription path was shows just a sequence of drugs or they combinations as they were prescribed; while the treatment path shows a sequence of states as derived from the state-transition model

Treatment path

According to methodology described in [5] medications were clustered into 6 groups (group B has 3 subgroups):

- Group A: Angiotensine converting enzyme inhibitors (ACEi) - ATC codes C09A and C09C
- Group B1: -blockers (BB) - ATC codes C07AA, C07AB
- Group B2: Diuretics - ATC codes C03AA, C03CA or C03D
- Group B3: Non-dihydropyridine Ca-channel blockers (NCCB) - ATC code C08D
- Group C: Dihydropyridine Ca-channel blockers (DCCB) - ATC code C08CA
- Group D: -blockers, hydralazines, clonidine - ATC codes C02CA, C02DB, C02AC

In our study we deviated slightly from standard ATC coding - combination drugs were coded as if a set of separate drugs was given. This deviation was useful to simplify the model-building algorithms and did not have impact on the analysis.

Based on prescription data states were generated, using the amount and the daily dose of prescribed drug to calculate duration of treatment with a particular drug group. In this calculation we assumed, that patient starts the medication on the same day as it was prescribed, that the dosage is unchanged throughout the time covered by a particular prescription (and the patient adheres to the dosage).

Prescription path

A prescription is a sequence of prescription instances. A prescription instance is a drug, or combination of drugs as prescribed on a particular day, disregarding dosage or quantities prescribed.

All methods were implemented using Cache and Ensemble (Intersystems Inc.), graphs were generated using GraphViz software (www.graphviz.org).

Results

A prescription path and a treatment path were created for each patient. Both paths are shown simultaneously and provide different views on what was done. E.g. if a drug containing a BB and diuretic was prescribed and at the same time a potassium-sparing diuretic was prescribed the prescription instance is shown as B1+B2+B2; while the treatment path will show B1+B2 combination. A particular prescription path can have several treatment path counterparts (Figure 1).

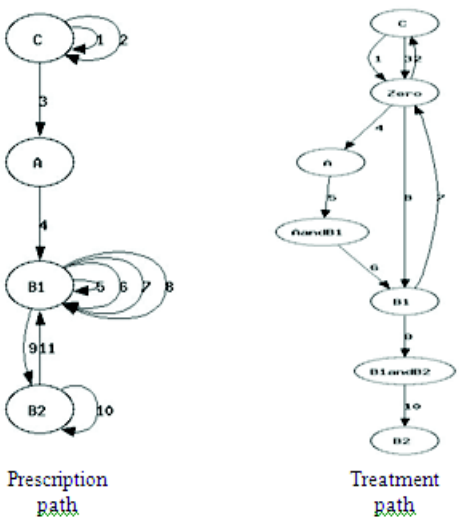


Figure 1 Prescription path vs. Treatment path

Alerts

Visualisation of the paths should be in most cases enough for an experienced GP to see any unusual patterns, however the results can be analysed in background and selecting only patients with an unusual pattern. Analysing treatment paths allows to create an additional class of alerts, taking into account more than just one state or one transition. More than just the preceding state (eventually the whole path) can be taken into account in launching an alert - e.g. transition from A to B1 as well as transition from B1 to A might be OK per se, but if this is repeated more than **n** (usually 2-4) times, it may be considered an unusual pattern and an alert should be launched.

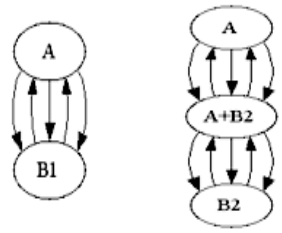


Figure 2 Cyclic patterns in treatment paths

Discussion

Guideline can be utilized at two levels in data analysis. First approach is to use the guideline to structure the problem space and then fit the data into the resulting of structure. Svátek and íha [8,9] used a guideline to lead a data-mining exercise in their analysis of guideline compliance in hospital environment. This idea is similar to what we use, however it was applied in a much better controlled environment of a large hospital. Others [10, 11, 12] focussed more on intentions and worked in a hospital, data rich setting. Major advantage of this approach is, that results are presented to the physician in the context of the guideline making it easier to recognise and appreciate the rationale behind the graphs and alerts.

Alternate approach is to use data mining techniques to raw data and then compare the results to a guideline. Possible method of creating similar results to ours is path or workflow mining (e.g. [14]). This approach can be exploited either to improve the guideline or to detect exceptions in clinical workflow.

Conclusion

Authors extended the scope of previous research [1] by adding visualizations for individual patients as well as new types of alerts taking into account more than just one transition or just one therapeutic state. This approach is to be validated by a clinical study in next future.

Acknowledgments

The study protocol was approved by the Human Research Ethics Committee of the University of South Australia (protocol P005-04) and undertaken with a Memorandum of Understanding with Lubims Pty. Ltd. Special thanks to Intersystems Inc. for providing licenses for Cache and Ensemble.

References

[1] Gadzhanova, S., Iankov, I.I., Warren, J.R., Stanek, J., Misan, G.M., Baig Z., Ponte, L.: Developing High-Specificity Antihypertensive Alerts by Therapeutic State Analysis of Electronic Prescribing Records. JAMIA, 2006 Oct 26 [epub. ahead of print]

[2] Sullivan, F., Wyatt, J.C. How computers help make efficient use of consultations. BMJ 2005;331:1010-1012.
[3] Western M, Dwan K, Makkai T, Del Mar C, Computerisation in Australian general practice. Aust Fam Phys 2003; 32(3): 1-6.
[4] Thiru K, Hassey A, Sullivan F, Systematic review of scope and quality of electronic patient record data in primary care. BMJ 2003; 326: 1070-80.
[5] Warren JR, Gadzhanova S, Stanek J, Iankov II, Understanding caseload and practice through analysis of therapeutic state transitions. Proc AMIA Annu Fall Symp 2005.
[6] Warren JR, Stanek J, Gadzhanova S, Iankov I, Misan G, Inferring 'therapeutic states' of patients from community electronic prescribing data. electronic Journal of Health Informatics 2006; 1(1): e5.
[7] American Diabetes Association, Hypertension management in adults with diabetes. Diabetes Care 2004; 27:S65-S67.
[8] Ríha A., Svátek V., Nmec P., Zvárová J.: Medical guideline as prior knowledge in electronic healthcare record mining. In: 3rd International Conference on Data Mining Methods and Databases for Engineering, Finance and Other Fields, 25-27 September 2002, Bologna, Italy.
[9] Svátek, V., Ríha, A., Peleška, A., Rauch, J.: Analysis of guideline compliance – a data mining approach. Praha 13.04.2004 – 16.04.2004. In: Kaiser, K, Miksch, S., Tu, S. W. (ed.). *Computer-Based Support for Clinical Guidelines and Protocols.* Amsterdam, IOS Press, 2004, s. 157–161. ISBN 1-58603-412-X.
[10] Advani, A.,Shahar, Y., Musen, M.A.: Medical Quality Assessment by Scoring Adherence to Guideline Intentions. 2001 Annual AMIA Conference, Washington, DC, Hanley and Belfus 2001.
[11] Chan, A.S., Coleman, R.W., Martins, S.B., Advani,A , Musen,M.A., Bosworth, H.B., Oddone, E.Z., Shlipak, M.G., Hoffman,B.B., Goldstein, M.K.: Evaluating Provider Adherence in a Trial of a Guideline-Based Decision Support System for Hypertension.MedInfo 2004, San Francisco, California, Sept 7 - 11. 2004.
[12] Marcos, M., Berger, G., van Harmelen, F., ten Teije, A., Roomans, H., Miksch, S: Using Critiquing for Improving Medical Protocols: Harder than It Seems. In: Quaglini, S., Barahona, P., and Andreassen, S. (Eds.): AIME 2001, LNAI 2101, Springer-Verlag Berlin Heidelberg 2001, pp.431–441
[13] Aigner, W, Miksch, S.: Supporting Protocol-Based Care in Medicine via Multiple Coordinated Views. Proceedings of the Second International Conference on Coordinated & Multiple Views in Exploratory Visualization (CMV'04), 2004.
[14] van der Aalst, W., Weijters, T., Maruster, L.: Workflow mining: discovering process models from event logs. IEEE Transactions on Knowledge and Data Engineering, 16 (9), September 2004.

Address for correspondence

Dr. Jan Stanek
University of South Australia
Advanced Computing Research Center
Mawson Lakes Blvd.
Mawson Lakes, SA 5095
Australia.
E-mail: jan.stanek@unisa.edu.au

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Efficiency and Safety of New Radiofrequency Identification System in Japanese Hospital

Yuichiro Saito[a], Takashi Hasegawa[b], Tetsuo Sakamaki[c]

[a]*Division of Cardiovascular Medicine, Gunma University Hospital, Maebashi, Japan*
[b]*Informatics Education Center , International University of Health and Welfare, Tokyo, Japan*
[c] *Medical Informatics and Decision Sciences,Gunma University Hospital, Maebashi, Japan*

Abstract and objective

Radiofrequency identification (RFID) uses radio-frequency tags attached to people or objects to provide identification, tracking, and security under the general heading of automatic identification. New RFID system (UHF band, 953 MHz) has been available since April, 2005 in Japan. We tested efficiency and safety of new RFID system in our hospital. Electric fields produced by new RFID had no significant effects on cardiac pacemakers, implantable cardioverter defibrillators (ICD), and other medical devices□@(ex. Electrocardiogram recorder, cardiac monitor, intra-aortic balloon pumping, infusion pump, and respirator) in our hospital. New radiofrequency tags seemed to provide extensive patient identification and to track capital equipments within our hospitals. Healthcare systems today are increasingly complex and interrelated processes, while new RFID technologies will provide opportunities for enhanced patient care and safety in Japanese hospital.

Keywords:

cardiac pacemaker, implantable cardioverter defibrillators, healthcare, patient safety

Introduction

Radiofrequency identification (RFID) has recently begun to receive increased interest in supply chain, in order to increase the efficiency and visibility of material and information flows. RFID may address to improve safety and increase in productivity. New RFID system (UHF band, 953 MHz) has been available since April, 2005 in Japan. However, there has been no attempt to determine whether electric fields produced by new RFID can influence cardiac pacemakers, implantable cardioverter defibrillators (ICD), and other medical devices in our hospital. We further examined to track new IC tags attached to people and medical equipments in hospitals.

Our data may provide the potential benefits, the area of applications, and the corresponding strategies of RFID in hospital environments.

Methods

We examined effects of RFID system on 7cardiac pacemakers and 5 ICDs in 0.18 % salt solution, similar electric condition to human body. The experimental conditions for interference induced by RFID were between a homogeneous electric field perpendicular to the area formed by the antenna and radio-frequency tags.

We next tested whether Electrocardiogram (ECG) recorder, ECG monitor, intra-aortic balloon pumping (IABP), infusion pump, and respirator in our hospital could work normally under new RFID system.

Furthermore, we evaluated to track new IC tags attached to people and medical equipments in our hospital.

Results

Electric fields produced by RFID had no significant effects on cardiac pacemakers. There was no differentiation between a unipolar and a bipolar system. The RFID systems did not interfere with ICDs.

ECG recorder, ECG monitor, IABP, infusion pump, and respirator were worked normally under operation of RFID systems.

Radiofrequency tags attached to people and capital equipments within our hospitals provide more extensive identification than traditional bar coding can.

Conclusion

RFID may be ultimately used for many of the functions currently carried out using bar coding if the cost of RFID comes down. Healthcare systems today are increasingly complex. RFID is a technology that will have a profound impact on effective and safe patient care in Japanese hospitals in near future.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Development of Hypertension Management Ontology for Guideline-based Clinical Decision Support System

JiHyun Kim[a], InSook Cho[b], EunJung Lee[a], JaeHo Lee[c], Yoon Kim[d]

[a] *Center for Interoperable EHR, Korea*
[b] *Department of Nursing, Inha University College of Medicine, Korea*
[c] *Department of Emergency Medicine, University of Ulsan College of Medicine, Korea*
[d] *Department of Health Policy and Management, Seoul National University College of Medicine, Korea*

Abstract

For knowledge representation of guideline based hypertension management, hypertension management ontology was developed. Ontology was adopted as a method for enabling to share and to reuse knowledge of a domain with consistency in this study. This work is an essential course in representing knowledge in a computerized method and enables to reuse knowledge in interested domain areas. On the basis of published guidelines including JNC 7, concepts related to hypertension management knowledge and relations between concepts were made and defined. Entities of concept were defined. This ontology includes high level 11 classes and about 300 concepts and it defines instances of each concept. This work will be use useful to build a tool in order to acquire specific domain knowledge and contribute to spreading knowledge sharing and standardized clinical guidelines.

Keywords:

ontology, hypertension management, computerizing clinical guideline, knowledge representation

Introduction

In Korea, the Government leads Electronic Medical Record (EMR) development for the 152 primary health centers, which is planned to be expanded to Electronic Health Record (EHR) system. As a part of this work, development of clinical decision support system (CDSS) is in the process of advancing. This study describes a part of constructing the guideline based decision support system for hypertension management in primary care setting and encoding hypertension management knowledge. This is essential for representing knowledge in a computerized method. It is for reusing knowledge in interested domain areas.

For knowledge sharing nationally, knowledge management method, such as ontology is main issue on the aspect of clinical decision support system and others. Ontology is a methodology to formalize a shared understanding of a domain and enables to share and reuse of knowledge consistently between software applications and humans.

Materials and methods

This ontology is primarily based on 12 published hypertension guidelines including JNC 7. Concepts related to hypertension management were defined at the modeling stage and defined concepts were classified according to semantic category. Relations and attributes of concepts were defined. We authored knowledge in Protégé 3.1.1 environment.

Results

About 300 concepts related to hypertension management were defined. A medication prescription, laboratory test, physical examination, and other related conditions were included in related concepts. Based on the concepts defined, high level classes which contained medication, clinical finding, test, patient education, problem, patient case, event, rule, therapy adjustment, temporal predicate, and eligibility criteria were made and defined. Each class included subordinate concepts of is-a relation. Then, attributes of concepts were defined. Completed ontology had 11 classes and about 300 concepts.

Conclusion

Through this approach, Creation of a tool to acquire particular domain knowledge is expected to be easier. Furthermore, the ontology of hypertension management for guideline based clinical decision support system is considered to contribute for sharing and disseminating knowledge and to spreading standardized clinical guidelines through reusing knowledge.

Address for correspondence

Center for Interoperable EHR, Annex Building, Seoul National University College of Medicine, 199-1 Dongsoong-Dong, Jongno-Gu, Seoul, 110-810, Korea, Kijii90@snu.ac.kr

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Secure Remote Access for Web Based Clinical Information System Using Policy Control of PCs and Healthcare PKI Authentication

Katsuya Tanaka[a], Mayumi Yoshida[b], Ryuichi Yamamoto[b]

[a] *Department of Planning, Information and Management, University of Tokyo Hospital*
[b] *Interfaculty Initiative in Information Studies, Graduate School of University of Tokyo*

Abstract and objective

This paper describes a robust method of secure remote access for Web based clinical information system using Health-care PKI authentication. It enables medical staffs to refer the medical data using his own PCs from their home or their mo-bile phones over the Internet. This system contributes to re-duce strain of medical staffs especially in such institutes where intensive care is carried 24h/365d. For this purpose, it must permit the remote access to the hospital information system from PCs in various conditions, it is necessary to establish a secure connection using VPN or SSL, and control the policy of client PCs for the prevention of computer virus effect etc., and at the last, confirm user authentication with strict identification using Healthcare PKI (HPKI).

Keywords:

healthcare PKI, clinical information system, secure remote access control, VPN

Introduction

In Japan, healthcare ICT is promoted by Japanese government and in this year, the government made "New IT Reform Strategy" and "The Action Plan 2006". It emphasizes the reengineering in the healthcare field using ICT. It includes some actual plan such as constructing secure and reliable network for regional and inter-regional cooperation, introduc-ing smart cards for healthcare, and development of Health-care PKI (HPKI) which is based on ISO 17090. Actually we started to use HPKI in 2004, as a demonstration experimental system in the University of Tokyo Hospital for secure remote access to the medical data. HPKI was used in this system for authentication purpose, with verification the attributes of the national licenses of healthcare professionals, such as medical doctors, registered nurse, etc. By using this system, a physician can access the medical data of his patients with his own PCs even in his home or his mobile phones.

Methods

In our hospital, medical staffs are able to access to medical records using web based system in addition to special client PCs. Mainly based on the web interface, the following parts were also developed.

1. Certification Authority (CA)

For the convenience of user registration operation, All CA function was implemented in a note PC. It is also used for registering HPKI certificate to USB token.

2. VPN gateway

The connection between the hospital network and a client PC is established using IPSec. We picked a VPN gateway, Cisco Systems/VPN 3005.

3. Policy Control System

At the establishment of VPN connection, the server side policy is downloaded to the client PC, and during the connection, client PC is under the predefined policy. We picked Zone lab/Integrity.

4. USB Token for HPKI

At the insertion of this USB token, HPKI certificate is copied to the certain repository in Windows, and on the removable of token, the repository is cleared.

5. Reverse Proxy Server

For the verification of the HPKI certificate, especially hcRole attribute, user access is only allowed via this SSL reverse proxy server. At the connection to an internal web server in hospital, the user certificate in client PC is pushed and verified. For the mobile access over the Internet not using HPKI, we developed CHTML converter gateway. The identification number of each mobile phone is stored at the registration in server side, and the access of unknown device is prohibited.

Results

Over 30 doctors participated in this experiment and the developed system is now in use. This system is effective to check the results of emergency laboratory tests, or radiological images on PACS system from remote place. This system requires various client software, such as IPSec client software, USB token driver and policy control client, so it is difficult to manage for the user to install or operation. We are now discussing for the improvement of this system, such as using smart card for the HPKI token, SSL-VPN technology for an easy management of client PCs.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Infobuttons: A Study of Usability

Lily Gutnik[a], Sarah Collins RN, BSN[b], Leanne M. Currie RN, DNSc[a,b], James J. Cimino MD[a], and Vimla L. Patel, PhD.DSc[a]

[a]*Department of Biomedical Informatics, Columbia University, USA*
[b]*School of Nursing, Columbia University, USA*

Abstract and objective

Studies of clinician's information needs while treating patients have shown that the resolution of these needs is often deferred or fails, which may lead to medical errors. The Infobutton Manager was developed to help improve the resolution of information needs by providing users with links to on-line health information resources. The aim of this study was to determine the usability of the Infobutton interface to resolve clinicians' information needs. We provided clinicians with typical case scenarios using a computerized order entry system (CPOE) and Infobuttons and asked them to verbalize their thought processes as they were using the CPOE. We video-recorded the computer screens as the users worked, conducted brief exit interviews, and analyzed these data. Results indicated that the participants found the resources provided by Infobutton helpful and easy to use.

Keywords:

usability testing, clinical information needs, information retrieval decision support

Introduction

Though technology for clinical care is created to facilitate the clinician's workload and enhance patient care, tools that are inadequately designed can actually have adverse effects [1]. Inadequate design can result from a disconnect between the design process and the needs of the end-users [2]. Infobutton is a tool designed to provide clinicians with access to on-line health information resources to quickly resolve their information needs [3]. This study explored the usability of a new user Infobutton interface which was designed to be easier to read, more concise, and have more consistent navigation.

Methods

We conducted the present laboratory study to seek in-depth feedback from clinicians about the design and usability of Infobuttons within a CPOE system. Each participant was given three typical scenarios to be interpreted using the CPOE system and Infobutton. The computer screen was video recorded using Morae™ software. Participants were asked to "think aloud," that is to verbalize their thoughts. At the end of the session, the researcher conducted a brief exit interview. A previously established coding framework was applied to the data to characterize information needs.

Results

Two nurses, one physician, and one physician's assistant took part in the study, yielding a total of 79 information needs. Twenty-three of these needs (29%) were related to drug information, 28 (35%) concerned institutional procedures or policies, and the remainder were related to patient care and the treatment plan. Forty-eight information needs (60%) were successfully resolved using Infobuttons, 15 were deferred, and 14 needs failed to be resolved. Fifty-one needs (65%) were from an external source (e.g., Micromedex) and 26 (33%) were from an internal source (e.g., the local intranet). Exit interviews revealed an overall satisfaction with the resources provided by and the usability of Infobuttons. Excess information making navigation difficult was one of the problems identified.

Conclusion

Although clinicians show a high occurrence of information needs as they treat their patients, many of these needs can be met with Infobuttons All participants agreed that the information resources provided by Infobutton were valuable and easy to understand. This study imparted us with insight on the questions clinicians need to have answered for effective decision-making, and accentuated the need to be mindful of information overload.

References

[1] Allen M, Currie LM, Bakken S, Patel VL, Cimino JJ. Heuristic evaluation of paper-based Web pages: a simplified inspection usability methodology. J Biomed Inform. 2006; 39(4):412-23.

[2] Kushniruk, AW, Patel, VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. J Biomed Inform. 2004; 37:56-76.

[3] Cimino JJ. Li J. Bakken S. Patel VL. Theoretical, empirical and practical approaches to resolving the unmet information needs of clinical information system users. Proc Ann AMIA Symp:170-4, 2002

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

The Application of A Clinical Data Warehouse to the Assessment of Drug-Warfarin Interactions

Qiyan Zhang[a*] PhD., Yasushi Matusmura[a] MD. PhD., Hiroshi Takeda[a] MD. PhD.

[a] *Department of Medical Informatics, Graduated School of Medicine, Osaka University*

Abstract and objective

Studies on drug-warfarin interactions in human subjects were typically based on single case reports or extrapolated in healthy volunteers. In this study, we proposed a method to apply an institutional clinical data warehouse (CDW) to address this issue in a real-world setting. A case-control study was conducted by using the CDW in Osaka University Hospital from 2000/01/01/-2005/12/31. We randomly selected the steady-state outpatients who were under warfarin mono-therapy as Group A and those who were on the pre-existing warfarin therapy but with only the added medication X as Group B. The difference between the PT-INR values on one-month interval in Group A and those before and after taking medication X for one month in Group B was compared. The warfarin-Allopurinol interaction was illustrated as one example. We identified 25 cases in Group A and 15 cases in Group B respectively. The difference of the PT-INR values on one-month interval between the two groups was not significant (P=0.394,Mann-Whitney test), indicating that no warfarin-Allopurinol interaction was present. This method can be used as an alternative approach to assess drug-warfarin interactions.

Keywords:

clinical data warehouse, warfarin, drug interaction, PT-INR

Introduction

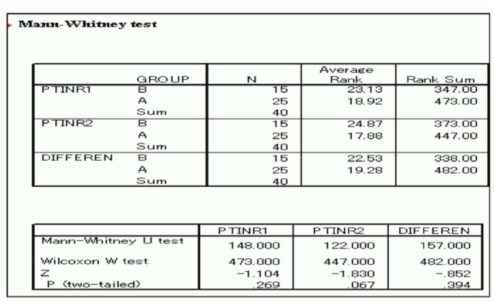
Warfarin is an effective and commonly used oral anticoagulant agent for the treatment and prevention of thromboembolism in a variety of conditions. The risk of major complication- hemorrhage may be increased when concomitant drug therapy is required. Previous studies on drug-warfarin interactions in human subjects were based on single case reports or extrapolated in healthy volunteers. Clinicians need to balance the therapeutic benefits with the bleeding risk through monitoring patient's PT-INR values during warfarin therapy. In this study, we proposed a method to apply an institutional clinical data warehouse (CDW) in a real-world setting to reduce certain practical and ethical problems faced by studies drawn from healthy volunteers and aimed to provide an alternative approach to assess drug-warfarin interactions.

Methods

The CDW in Osaka University Hospital was built in 1995 as a subject-oriented database. The outpatient's prescription data and PT-INR results were picked up from 2000/1/1/-2005/12/31 through Business Objects 6.5.1. Data were processed via Microsoft Access 2003 and statistic analyses were performed via SPSS 11.0 Japanese Version for Windows. A case-control study was conducted. The steady-state outpatients who took warfarin consecutively for at least one month were considered as the eligible study subjects. We randomly selected those who were under warfarin mono-therapy as Group A and those who were on the pre-existing warfarin therapy but with only the added medication X as Group B. Then we compared the difference between the PT-INR values with one-month interval in Group A and those before and after taking medication X for one month in Group B. The warfarin-Allopurinol interaction was demonstrated as one example.

Results

For the chosen example, we identified 25 cases in Group A and 15 cases in Group B respectively. The difference of the PT-INR values on one-month interval between the two groups was not statistically significant (P=0.394,Mann-Whitney test), indicating that no warfarin-Allopurinol interaction was present. **Figure1**



Mann-Whitney test

	GROUP	N	Average Rank	Rank Sum
PTINR1	B	15	23.13	347.00
	A	25	18.92	473.00
	Sum	40		
PTINR2	B	15	24.87	373.00
	A	25	17.88	447.00
	Sum	40		
DIFFEREN	B	15	22.53	338.00
	A	25	19.28	482.00
	Sum	40		

	PTINR1	PTINR2	DIFFEREN
Mann-Whitney U test	148.000	122.000	157.000
Wilcoxon W test	473.000	447.000	482.000
Z	-1.104	-1.830	-.852
P (two-tailed)	.269	.067	.394

Figure 1 - Mann-Whitney Test

Conclusion

This method can be used as an alternative approach to assess drug-warfarin interactions.

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

Implementation of An Integrated Network for Health Research in Quebec.

Allen Huang[a], Jacques Lemieux[b], Jean-Claude Bouchard[b], Michel Bourque[b], Robyn Tamblyn[b]

[a] *McGill University Health Centre, Montreal, Canada*
[b] *Clinical and Health Informatics Research Group, McGill University, Montreal, Canada*

Abstract

Health data warehouses represent a valuable resource for health research. We have developed an infrastructure capable of providing health researchers in the province of Quebec with a toolkit to access the clinical data warehouses contained in the major academic health centres and the provincial health administrative systems. This demonstration will highlight the components that allowed the successful implementation of an integrated network to accomplish this task. Acceleration of the pace and increases in the volume and quality of health research within the province, in other jurisdictions and possibly world-wide is now an attainable goal.

Keywords:

health databases, information management, patient data privacy, database management systems, computing methodologies, health services research

Introduction

Health research done in the conventional, paper-record environment is tedious, expensive and reliant on data with varying quality. Large repositories of health data are invaluable resources to researchers and planners in health care. In the province of Québec, Canada, these repositories are housed in clinical data warehouses within the large teaching hospitals and the administrative data warehouse at the Régie de l'assurance maladie du Québec (RAMQ) – the provincial health services payer. The Infostructure de Recherche Intégrée en Santé (IRIS) - Québec project is a Canada Foundation for Innovation funded initiative to construct an integrated network for health research in the province. Its goal was to create secure access to these data warehouses, enable the linkage of patient records through the use of a provincial Master Person Identifier (MPI) and ensure that resultant datasets returned to researchers conform to privacy standards.

Methods

The IRIS-Québec architecture is a distributed, federated data warehouse model. The RAMQ already manages the MPI. The research warehouses of clinical data were constructed to ensure the highest standards of data quality. People: Researchers wishing to access these data warehouses no longer have to queue for specialized data analysts and programmers. The researchers' toolkit is a web-based, user-friendly interface that drives a powerful system with the following functions: selection of variables from the extensive data dictionary across the warehouses, building of complex queries using logical operators, a temporal relation tool to define time-dependencies of variables, a crosstabs manager, and a data extraction manager. Previously onerous and lengthy authorization steps have been streamlined into an electronic approval process which is reliable, timely and track able. Privacy: A novel "inference controller" was developed to ensure individual data privacy. This software computes the probability of the presence of an unique, potentially re-identifying profile when multiple databases are linked. The researcher's toolkit allows the dynamic modification of data query parameters in order to achieve the desired data precision without violating privacy rules. Processes: Data sharing agreements were established with each partner institution. Software agents were installed at each data source to manage queries, connectivity, linkages, and data flows. Each agent contains a copy of the inference controller engine. Performance: is maintained by de-coupling the phases of a research project from the raw data extractions. The creation of the complete cohort data cube is triggered only on submission of the finalized project profile. We will demonstrate the operation of the Toolkit through sample queries.

Conclusion

The architecture of our system can potentially be extended to other biomedical information sources such as genomics, proteomics and geneology databases and create new capacity for health and bio-medical research. This architecture can also be potentially replicated in other jurisdictions. Scaling such a system to the international level can result in enabling health and bio-medical research in the global community.

Address for correspondence

Allen Huang, MDCM
McGill University Health Centre
687 Pine Avenue W, Room M8.12
Montreal, Quebec
Canada H3A 1A1

MEDINFO 2007
K. Kuhn et al. (Eds)
IOS Press, 2007

OpenECG: Promoting Interoperability Through the Consistent Implementation of the SCP-ECG Standard in Electrocardiography

Catherine Chronaki[a], Franco Chiarugi[a], Ronald Fischer[b]

[a]*Institute of Computer Science, FORTH, Heraklion, Crete*
[b]*Department of Biometrics, MHH, Hanover, Germany*

Abstract and objective

The OpenECG Network (www.openecg.net) has been created to promote interoperability in electrocardiography with tutorials, specifications, open source tools, data sets, converters, and interoperability testing. ECG vendors, members of professional organizations, researchers, and other stakeholders participate in the OpenECG network to exchange views and receive assistance in implementation. In 2006, members are more than 700 individuals from 58 countries. A specific focus area for OpenECG that concerns diagnostic quality resting electrocardiograms (ECGs) is SCP-ECG, the European standard (EN1064:2005). An online interoperability testing service assists members in consistently implementing SCP-ECG and effortlessly integrating electrocardiographs with eHealth systems. OpenECG is a case of best practice in interoperability that should be followed by medical devices and sensors for effective personalized health monitoring.

Keywords:
interoperability, standards, telemedicine, eHealth services, medical devices, electrocardiography

Methodology

Electrocardiography is the most frequently applied non-invasive examination for early detection of heart disease, a leading cause of morbidity and mortality in western countries. It is estimated that more than 100 million ECGs are recorded annually in Western Europe. The ECG allows early detection and follow-up of heart disease, but today the operation of most ECG devices is still based on proprietary protocols and file formats.

SCP-ECG is a standard communication protocol that specifies the interchange format and a messaging procedure for ECG equipment-to-computer communication and for retrieval of ECG records from the computer to the ECG equipment (if needed). Since March 2005, SCP-ECG (EN1064:2005) is the European standard for high quality diagnostic ECG exchange and if consistently implemented ensures interoperability. Nevertheless, it is rather difficult for integrators to implement SCP-ECG correctly and there are variations in implementations, which can be a barrier to interoperability. Although a number of ECG device manufacturers and integrators have implemented the SCP-ECG standard, most implementations are not fully accurate. This is due partly to misconceptions and partly to the lack of widely publicized conformance levels and IHE-like integration statements that exist for modalities in radiology.

In 2003, OpenECG established an online conformance testing service to support the OpenECG community at large in implementing interoperable eHealth systems with SCP-ECG support. A member may submit an ECG file in an alleged SCP-ECG format and receive a list of errors and warnings. If no errors are detected, the submitter may request a certificate that is granted after thorough manual review of the file.

Results

The SCP-ECG conformance testing service has been used extensively by members and in many occasions an interoperable solution was achieved with support from the help desk. In the period 2003-2006, more than 1700 ECGs were submitted for conformance testing by members in more than 20 counties worldwide. Leading is Italy with 11 members, who have submitted 37.3% of the tests. After Italy, most ECGs have been submitted by Greece (17.74%) and Hungary (11.89%). ECG devices and eHealth services have been tested, improved, and validated using online tools and support from the OpenECG helpdesk. In 2005, a web service variant of the conformance testing service was integrated to the ECG viewer that won the first prize in the OpenECG programming contest. After certain limitations of the software were identified and amended, a new version of the ECG viewer was released and is currently available at the OpenECG open source repository.

Conclusions

The OpenECG network promotes best practice in interoperability for ECGs. Innovative eHealth services capable of managing personal wellness profiles call for plug-interoperability of medical devices, which is an issue of patient safety, key to advancing quality and cultivating consumer trust in the next generation of ambient intelligent working and living environments.

Subject Index

This page intentionally left blank

Author Index

This page intentionally left blank

This page intentionally left blank

This page intentionally left blank